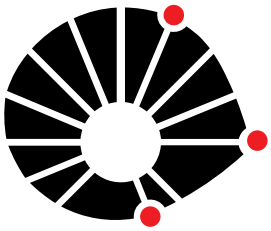


UNICAMP

UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística e Computação Científica

JACKELINE DEL CARMEN HUACCHA NEYRA

Developing an interior-point algorithm for the radiotherapy planning problem in a fuzzy approach background

Desenvolvendo um algoritmo de pontos interiores para o problema de planejamento de radioterapia em um contexto de abordagem difusa

Campinas

2021

Jackeline del Carmen Huaccha Neyra

Developing an interior-point algorithm for the radiotherapy planning problem in a fuzzy approach background

Desenvolvendo um algoritmo de pontos interiores para o problema de planejamento de radioterapia em um contexto de abordagem difusa

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutora em Matemática Aplicada.

Thesis presented to the Institute of Mathematics, Statistics and Scientific Computing of the University of Campinas in partial fulfillment of the requirements for the degree of Doctor in Applied Mathematics.

Supervisor: Aurelio Ribeiro Leite de Oliveira

Este trabalho corresponde à versão final da Tese defendida pela aluna Jackeline del Carmen Huaccha Neyra e orientada pelo Prof. Dr. Aurelio Ribeiro Leite de Oliveira.

Campinas

2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

H86d

Huaccha Neyra, Jackeline del Carmen, 1993-
Developing an interior-point algorithm for the radiotherapy planning problem in a fuzzy approach background / Jackeline del Carmen Huaccha Neyra. – Campinas, SP : [s.n.], 2021.

Orientador: Aurelio Ribeiro Leite de Oliveira.
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Métodos de pontos interiores. 2. Radioterapia. 3. Otimização fuzzy. I. Oliveira, Aurelio Ribeiro Leite de, 1962-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Desenvolvendo um algoritmo de ponto interior para o problema de planejamento de radioterapia em um contexto de abordagem difusa
Palavras-chave em inglês:
Interior-point methods
Radiotherapy
Fuzzy optimization
Área de concentração: Matemática Aplicada
Titulação: Doutora em Matemática Aplicada
Banca examinadora:
Aurelio Ribeiro Leite de Oliveira [Orientador]
Estevão Esmi Laureano
Daniela Renata Cantane
Weldon Alexander Lodwick
Carla Taviane Lucke da Silva Ghidini
Data de defesa: 08-09-2021
Programa de Pós-Graduação: Matemática Aplicada

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0000-0003-2162-1917
- Currículo Lattes do autor: http://lattes.cnpq.br/6121255962590593

Tese de Doutorado defendida em 08 de setembro de 2021 e aprovada pela banca examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). AURELIO RIBEIRO LEITE DE OLIVEIRA

Prof(a). Dr(a). ESTEVÃO ESMI LAUREANO

Prof(a). Dr(a). DANIELA RENATA CANTANE

Prof(a). Dr(a). WELDON ALEXANDER LODWICK

Prof(a). Dr(a). CARLA TAVIANE LUCKE DA SILVA GHIDINI

A Ata da Defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria de Pós-Graduação do Instituto de Matemática, Estatística e Computação Científica.

Este trabalho é dedicado aos meus pais, Pedro e Adelina,
todo seu esforço e sacrifício valeu a pena.

Acknowledgements

Agradeço aos meus pais, Pedro e Adelina, que, desde que souberam que fui aceita no doutorado, compartilharam minha alegria e sacrifícios apoiando cada uma das minhas decisões e compreendendo que não era possível estar presente em momentos difíceis e datas importantes. Aos meus irmãos, Stefhannie e Gabriel, com quem cada semana tentamos ter uma vídeo-chamada para desabafar e fofocar.

Agradeço ao meu companheiro, amigo, esposo, Heraclio López, pela motivação e força nos momentos em que pensei que não conseguiria.

Agradeço ao Professor Dr. Aurelio Ribeiro Leite de Oliveira, pela confiança, orientação, comprometimento e dedicação proporcionada no processo de doutorado.

Agradeço a cada um dos funcionários do IMECC que também contribuiram no processo.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Resumo

O problema de distribuição de doses no processo de planejamento de radioterapia consiste em otimizar a dosagem total de radiação para o tratamento de um paciente com câncer. O objetivo principal é atacar o tumor enviando tanta radiação quanto seja possível, evitando danos severos aos tecidos e órgãos que estão próximos do tumor. Para isso usamos a técnica de radioterapia de intensidade modulada (IMRT) que nos permite emitir fluxo de radiação não uniforme obtendo um melhor ajuste das doses de radiação para as células tumorais e um modelo de programação linear para formular este problema, surgindo múltiplas opções de escolha de funções objetivo como por exemplo maximizar a dose para o tumor, minimizar as doses para os tecidos saudáveis ou ambas. Além disso, na prática, o oncologista pode variar a dose de acordo com a sua experiência tornando difícil garantir uma quantidade exata de radiação. Assim, neste trabalho, as doses são consideradas como números fuzzy pois estes conjuntos descrevem melhor a imprecisão dos valores de doses atribuindo graus de pertinência onde o grau de valor 1 representará uma dose completamente aceitável para o oncologista. E, usamos a função surpresa, que pode ser entendida como uma penalidade contra a violação de cada restrição (ou seja, quando o fluxo de radiação emitido não é suficiente para o oncologista), para converter estas restrições fuzzy numa função não-linear que deve ser minimizada. Nesta tese apresentamos um modelo matemático com função objetivo, definida a partir da função surpresa para cada restrição, não-linear e convexa e restrições lineares limitando a quantidade de fluxo de radiação. Devido às caraterísticas da função objetivo, escolhemos os métodos de Pontos Interiores que, apesar de ser desenvolvidos para problemas lineares, funcionam bem quando as restrições são mantidas lineares. Definimos a função surpresa a partir da função de pertinência do número fuzzy, porém existem, na literatura, uma grande variedade de números fuzzy, sendo os números fuzzy triangulares e trapezoidais os mais clássicos. Intuitivamente pode-se pensar que tratar as doses como números fuzzy trapezoidais é melhor que os números triangulares devido a que a quantidade de valores que têm grau 1 é maior, isto é, o conjunto de valores aceitáveis para o oncologista é maior, fazendo com que o algoritmo permaneça no mesmo ponto. Levando em consideração estas situações indesejáveis, tomamos os valores de dose como números fuzzy triangulares, e adicionamos restrições lineares a fim de obter boa qualidade de soluções e bom desempenho do algoritmo proposto que foi desenvolvido usando um método de Pontos-Interiores Primal-Dual especialmente adaptado para resolver este problema. Nosso algoritmo foi implementado usando MATLAB, pois usamos algumas das funções já definidas no software como a função de mínimos quadrados para obter o ponto inicial baseado na dose do tumor e iterar buscando a dose mínima nos outros tecidos; e testado em problemas reais de grande porte onde o tumor está localizado na área da cabeça e pescoço, ressaltando que o planejamento das doses e valores mínimos aceitáveis é de acordo à equipe médica. Apresentamos os resultados numéricos em tabelas onde

mostramos os valores das doses depositada em cada região, de cada paciente. Também apresentamos os histogramas Dose × Volume (%) que mostram a porcentagem de tecido que recebe certa quantidade de radiação. Ao analizar as tabelas e os histogramas observamos que a dose depositada no tumor é mais do que a suficiente de acordo com a equipe médica que forneceu o conjunto de dados; e nas outras regiões do paciente está dentro dos níveis permitidos para garantir a funcionalidade dos órgãos. De modo geral, podemos concluir que a abordagem desenvolvida fornece soluções favoráveis, com o conjunto de dados usado, para o problema de distribuição de dose.

Palavras-chave: Método de Pontos interiores. Radioterapia. Otimização Fuzzy.

Abstract

The dose distribution problem in the Radiation Therapy planning consists in optimize the total radiation dosage delivered into the patient with cancer. Its main aim is to attack the tumor delivering as much radiation as possible, while avoiding severe damage the nearby tissues and organs. For that we use the intensity modulated radiation therapy (IMRT) technique that allows us to deliver non-uniform radiation flow obtaining a best adjustment of radiation doses to the tumor cells and a linear programming model to formulate this problem, arising multiple options for choosing objective functions as to maximize the dose delivered to the tumor, to minimize the dose delivered to the healthy tissue or both of them. However, in practice, the oncologist can vary the dose according to his/her expertise making it difficult to guarantee an exact quantity of radiation. Therefore in this work, the doses are considered as fuzzy numbers since these sets have a best description on an imprecision for the value dose attributing membership degrees where the degree 1 represents a dose totally acceptable by the oncologist. And, we use the surprise function, that can be understood as a penalty for the violation of each constraint (that is, when the delivered radiation flow is not enough for the oncologist), to translate these fuzzy constraints in a non-linear function that must be minimized. In this thesis we present a mathematical model with non-linear and convex objective function, defined from the surprise function for each constraint, and linear constraints for bounding the radiation flow. Due to the objective function characteristics, we choose the Interior-Point Method that, despite being designed for linear problems, it works well when the constraints are maintained linear. We define a surprise function from membership function of the fuzzy number, but there exists, in the literature, a great variety of fuzzy numbers, being the triangular and trapezoidal fuzzy numbers the most classical. Intuitively, one might think that treating the dose as a trapezoidal fuzzy number is better than a triangular number because the amount of values that have degree 1 is greater, that is, the set of values acceptable for the oncologist is greater, making the algorithm staying in the same point. Taking account these undesirable situations, we consider the dose values as triangular fuzzy numbers, and we add linear constraints in order to obtain a good quality of solutions and a good performance of the proposed algorithm that was designed using a specially tailored Primal-Dual Interior-Point Method to solve this problem. Our algorithm is implemented in MATLAB, because we use some of the functions already defined in this software as the least-square method to compute the initial point based on the tumor dose and then iterating to the minimal dose for the other tissues; and tested in real world large-scale problems in which the tumor is localized in the Head and Neck area, highlighting the planning of dose and acceptable minimal values is according to the medical time. We present the numerical results in tables where we show that the values of the dose delivered in each region, of each patient. Also, we present the histograms Dose $\times$ Volume (%) that

show the percentage of tissue that received certain quantity of radiation. When analyzing the tables and histograms we observed that the dose delivered in the tumor is more than the enough value, according to the medical time that provides the dataset; and in the other regions of the patients the dose is inside the allowed values in order to guarantee the functionality of them. In general, we can conclude that the developed approach provides favorable solutions, with the dataset used, for the dose distribution problem.

Keywords: Interior-Point Methods. Radiation Therapy. Fuzzy Optimization.

List of Figures

List of Tables

List of symbols

$\mathbb{R}^n$	set of $x = (x_1, \ldots, x_n)$ such that $x_i \in \mathbb{R}$, for all $i = 1, \ldots, n$
$\mathbb{R}^{m \times n}$	set of real matrices with dimension $m \times n$
S	set of feasible points ($S \subset \mathbb{R}^n$)
f	real function, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$
F	vector function, $F : \mathbb{R}^m \longrightarrow \mathbb{R}^m$
∇f	Gradient of the function f
H	Hessian matrix of the function f
$J(\cdot)$	Jacobian of the function F
$\mathcal{L}$	Lagrangean function
X	diagonal matrix, $X = diag(x)$
e	n-dimensional vector of 1's
$\mu(\cdot)$	membership function, $\mu : \mathbb{R} \longrightarrow [0, 1]$
$\mathcal{S}(\cdot)$	surprise function
$\tilde{b}$	vector of fuzzy numbers

List of Algorithms

Contents

Introduction

Linear Programming (LP) is an area of the applied mathematics that solves real-world problems formulating mathematical models and using computational methods to find solutions according to the goals and satisfying a set of constraints. In mathematical terms, LP problems consist of finding solutions to satisfy a finite system of linear equations and inequalities that minimize a linear objective function. One of the first uses of the LP was during the World War II, when the American army realized that the traditional methods and procedures of allocation and expense planning weren't yielding much profit, then decided create an office of operations research in order to formulate and implement models that allow reduced costs and increase the enemy losses. This kind of research required the collaboration of many scientists from various fields, for example: mathematicians, economists, statisticians, engineering, etc. who used the Linear Programming as a tool [15].

George B. Dantzig, in 1947 developed and presented a solution method to solve LP problems, the Simplex method [11], the most used method. Despite the Simplex method presented, in terms of computational performance, being superior to any other known procedure for solving LP problems, mathematicians wanted to find a procedure with the polynomial property. After many decades, in 1984, Karmarkar [17] was be able to develop a polynomial and efficient method, called the Interior-point method.

There are, in the Interior-point method, exist different approaches to proceed towards the solution: primal, dual and primal-dual methods. We will focus on the last one since it has excellent theoretical properties and practical performance, Stephen J. Wright (1997) [28]. Thanks to these characteristics, primal-dual interior-point methods have been used over the years to solve a wide variety of complex real world problems. One of these type of problems, that affects all countries, is to fight cancer.

The World Health Organization has stated that cancer is the second leading cause of death globally, in 2018, with an estimated of 9.6 million deaths[1] and due to life style, financial situation, even health systems, among others, this number tends to increase. However, when the cancer is identified early, it is possible to use effective treatment leading to a high probability of surviving.

Nowadays there are several options of cancer treatment, many patients are treated with the Radiation Therapy that basically consists of delivering radiation to kill cancer cells and shrink tumors. Depending on the localization of the tumor the radiation therapy can be internal or external.

[1] Site: https://www.who.int/health-topics/cancer#tab=tab_1

The internal radiation therapy, also called brachytherapy, is a method in which the radioactive sources are allocated inside the patient. It is generally used when a small area of the patient needs a high dose of radiation. The external radiation therapy, the focus of this work, also called external beam therapy or teletherapy, is a method for delivering radiation to a patient's tumor where no radioactive sources are placed inside the patient's body[2]. This radiation is generated outside the patient (usually by a linear accelerator, see Figure 3) and are targeted at the tumor site.

The objective of the radiation therapy is to kill the tumor delivering as much radiation as possible, avoiding damage to the nearby healthy organs and tissues. Thus. plan for a treatment of distribution of radiation must be created. The problem arises when the treatment distributes little radiation to the tumor or too much radiation to the healthy organs compromising their functionality and reducing the quality of the patient's life. Therefore it becomes necessary to develop a radiation distribution plan focused on shrinking the tumor (or in the best cases, removing it) and at the same time insuring that the various organs and healthy tissues surrounding the tumor do not receive high radiation.

This, the Intensity-modulated radiation therapy (IMRT), proposed by Brahme in 1988 [4], a technique, which delivers non-uniform radiation flow (called intensity modulated beams (IMB)) is used to achieve this allowing a best adjustment of radiation doses to the tumor cells. Rehman et al. (2018) [26] presented a brief review of the advances and limitations of the IMRT techniques, highlighting its ability of to model precise dose to tumors of complex concave shapes which facilitates the sparing of surrounding organs.

There exists two ways of planning external radiation therapy. A forward planning, where the parameters of the beam are chosen by trial and error until an acceptable treatment plan be found, that is, until achieve an acceptable distribution of dose. And inverse planning, that uses computerized optimization methods to generate a suitable plan, where the parameter or initial data is the absorbed dose and the compute of deliver dose and selected beams (the most difficult part of the process) are determined with the help of operational research techniques.

Fortunately, this kind of problems can be formulated as an optimisation problem where the radiation delivered into the tumor and the surrounding structures, satisfies a dosage provided by the radiation oncologist. That is, the linear programming can be used to formulate the Radiation Therapy problem.

The first optimization model was proposed by Bahr et al. (1968) [1]. They used the linear programming to formally express the mathematical problem, stating:

- A set of variables which will be non-negative in a feasible solution. The object is to

[2] https://www.radiologyinfo.org/en/info.cfm?pg=ebt

find the values of each of these variables such that its values produce the optimum solution.

- A set of constraints which bound the feasible solutions. These constraints must be expressed as linear relationships among the variables.
- A specification of bounds which must be complete, i.e., any set of values which fall within the limits of the constraint equations is a feasible acceptable solution.
- A definition of "optimum solution" which can be expressed as a linear relationship among the variables.

Thenceforth, many researches developed and experimented new and innovative techniques for treating cancer patients with radiation using linear models. Shepard et al. (1999) [22] presented several approaches to optimize treatment plans in radiation therapy, and explored the advantages and disadvantages of a number of formulations which considered a variety of objective functions and constraints.

As it was expected, linear programming models do not always find feasible solutions. In order to aid in the design of radiotherapy plans, to solve the infeasibility problems, Holder (2003) [14] introduced a new linear programming model that incorporates elastic constraints. He presented two theorems that guaranteed the existence of feasible primal and dual solution for his formulation and solved it with a path following interior-point method.

Ehrgott et al. (2010) [13] presented a survey of the use of mathematical models and optimization methods. They discussed the advantages and disadvantages of each model classifying them in feasible problems, linear, non-linear, mixed-integer and multiple objective models.

Breedveld et al. (2012) [8] constructed the i-Cycle algorithm to solve the dose distribution problem using multi-criterial optimization of beam angles and IMRT profiles resulting in significant improvements in treatment plan quality. In order to achieve highest efficiency and different applications, designed an implementation with three extensions to the standard interior-point method that describe a starting approach and, since the primal-dual interior-point method is considered as the most robust method to solve large-scale problems, Breedveld et al. (2017) [5], the use of non-convex cost-function and steplength control.

There is no genral agreement in the literature on the objective function. This function should minimize total radiation, maximize minimum tumor dosage, minimize radiation to critical structures, etc.

It is difficult in practice to ensure that with an exact quantity of dose since the oncologist can vary the dose distribution according to his/her expertise and tolerances

of each region. Thus, Lodwick et al. (2001)[19] presented three approaches to use fuzzy optimization in the radiation therapy problems where the fuzziness appears on the right-hand side. One of these approaches considers the dosage values as fuzzy numbers and uses the theory of surprise functions, developed by Neumaier (2003) [20], to translate the fuzzy constraints into non-linear programming problem.

Lodwick and Bachman (2005)[18] presented three approaches to uncertainty problems. They solved these problems using surprise methods developed by Tanaka et al. (1974) [25] and Zimmermann (1976) [30], for formulations with fuzzy inequalities (called as *soft constraints*); and Neumaier (2003) [20] compare the solutions with those formulations with the right-hand side being a fuzzy number. Moreover, Jamison and Lodwick (2001) [16], developed models when at least one parameter is a possibility distribution. They used MATLAB's optimization toolbox showing that is possible to solve efficiently large fuzzy optimization problem. On their experiments they highlighted that the surprise approach is robust for handling large problems but with memory restrictions.

Thus, unlike to Lodwick and Bachman [18], we propose to solve the large-scale fuzzy optimization problem applying a tailored primal-dual interior-point method, with the membership function defined from the fuzzy number of the right-side, because taking a possibility measure (as [18]) small dose values delivered on the tumor voxels will be highly possible moving away from the desired value. Due to the surprise function, the formulation has non-linear and convex objective function. Since the numerical experience indicates that interior-point methods, (despite being designing for linear problems, they work well when the constraints are maintained linear) are strong on large-scale applications and have been used successfully for non-linear optimization [21], we will develop an extension of the method in order to design an algorithm that will solve the dose distribution problem where the measure of violation of the dose constraints is translate using the surprise function and minimizing it.

The Chapter 1 presents the basic concepts and theorems of the linear programming, describes the interior-point methods and also its modifications when the objective function is non-linear. The Chapter 2 presents a briefly review about the fuzzy theory enunciating some definitions and properties of fuzzy numbers and the connection between membership functions and surprise functions. These concepts helps us to understand the fuzzy optimization formulation. In order to use the fuzzy approach on radiation therapy problems, in the Chapter 3 we begin describing the process of radiation therapy planning and then present a mathematical formulation of the problem. Next, in the Chapter 4 we present the radiation therapy formulation with the surprise function, then we develop the primal-dual interior-point method considering the properties of the problem, adding suitable constraints and establishing parameters for the implementation of the proposed method. We use the MATLAB software to implement our algorithm and in the Chapter 5

we present the numerical experiments where we test our algorithm with the large-scale dataset [6] of Head-and-Neck cancer patients. Finally, in the Chapter 6, we give the conclusions concerning the work and the future proposals.

1 Linear Programming

Linear programming arises as a proposal to solve (or analyze) real-world problems transforming these decision problems into mathematical models together with suitable algorithms.

1.1 Basic concepts and theorems

A classical formulation of a linear programming (LP) problem is

$$\begin{cases} \min & c^T x \\ s.t & Ax \leqslant b \\ & x \geqslant 0, \end{cases} \tag{1.1}$$

where $c \in \mathbb{R}^n$ is the vector of costs, $A \in \mathbb{R}^{m \times n}$ is the matrix of constraints, $b \in \mathbb{R}^m$ is the vector of independent terms and $x \in \mathbb{R}^n$ vector of variables.

The theory existing to describe and analyze linear optimization problems is often developed for the standard formulation, which is easy to obtain adding non-negative variables called *slack variables* to inequality constraints [2],

$$(P) \begin{cases} \min & c^T x \\ s.t & Ax = b, \\ & x \geqslant 0. \end{cases} \tag{1.2}$$

with $rank(A) = m < n$.

Definition 1.1. *$x \in \mathbb{R}^n$ is said feasible point of (P) if satisfies all constraints, that is, $Ax = b,\ x \geqslant 0$.*

Definition 1.2. *The set of feasible points $S = \{x \in \mathbb{R}^n : Ax = b,\ x \geqslant 0\}$ is called feasible set (or feasible region). This set is convex [2].*

Thus, solving a LP problem means to find x^* feasible ($x^* \in S$) such that the value of a preestablished objective function ($f(x) = c^T x$) be minimal, that is, $f(x^*) \leqslant f(x)$, for all $x \in S$. This point, x^*, is called an *optimal solution of (P)*.

Now, let us define the dual problem for (1.2),

$$\begin{cases} \max & b^T y \\ s.t & A^T y \leqslant c, \\ & y \text{ free}. \end{cases}$$

Adding the non-negative slack dual variable $z \in \mathbb{R}^n$, we get the standard dual formulation

$$(D) \begin{cases} \max & b^T y \\ s.t & A^T y + z = c, \\ & z \geqslant 0, \ y \text{ free.} \end{cases} \tag{1.3}$$

Thus, formulation (1.2) is called *primal problem*. As any LP problem can be written in the standard form, there exists the dual for all LP problems [2].

Analogously to the primal problem, the point (y, z) is a feasible solution of (D) if $A^T y + z = c$ and $z \geqslant 0$. And, consequently (y^*, z^*) is an optimal solution if is feasible and $b^T y^* \geqslant b^T y$ for all feasible (y, z).

The main properties and relations between (P) and (D) are [2], [28]:

Lemma 1.1. ***Weak duality property:*** *If x and (y, z) are feasible solutions for (1.2) and (1.3), respectively, then $c^T x \geqslant b^T y$.*

Note that, from Lemma 1.1, a feasible dual point provides a lower bound for the primal problem and a feasible primal point provides an upper bound for the dual one.

Corollary 1.1. ***Optimality property:*** *If x^* and (y^*, z^*) are feasible solutions for (1.2) and (1.3), respectively, and $c^T x^* = b^T y^*$, then x^* and (y^*, z^*) are optimal solutions for (1.2) and (1.3), respectively.*

Corollary 1.2. ***Unboundedness property:*** *If the primal (dual) problem has an unbounded solution, then the dual (primal) problem is infeasible.*

Corollary 1.3. ***Strong Duality Property:*** *If the primal (dual) problem has a finite optimal solution, x^*, then so does the dual (primal) problem, (y^*, z^*), and the objective function values are equal, $c^T x^* = b^T y^*$.*

Theorem 1.1. ***Complementary Slackness Property*** *Let x and (y, z) be feasible solutions for (P) and (D), respectively. Thus, x and (y, z) are optimal solutions if only if for all $i = 1, \ldots, n$*

1. $x_i > 0 \Rightarrow z_i = 0$,
2. $z_i > 0 \Rightarrow x_i = 0$.

Observe that, Lemma 1.1 leads us to an useful concept: the *duality gap*, also known as *gap*, is defined as the difference between the primal-objective-function and the dual-objective-function, that is $gap = c^T x - b^T y$ which is $\geqslant 0$ for feasible primal and dual points. It is easy to show that $gap = x^T z$ for feasible points [2]. Therefore, from Corollary 1.1, in the optimal solutions we have $gap = 0$ [21].

If $x > 0$, x is called *interior point* for the primal problem; if $z > 0$, (y, z) is an interior point for the dual problem.

1.1.1 Optimality Conditions

The optimality conditions for a LP problem are obtained as a particular case of the Karush-Kuhn-Tucker conditions (also known as KKT-conditions).

Theorem 1.2. ***Primal Optimality Conditions*** *The point $x \in \mathbb{R}^n$ is a solution of (P) if only if there exist vectors $y \in \mathbb{R}^m$ and $z \in \mathbb{R}^n$ such that*

$$\begin{aligned} Ax &= b, && \textit{Primal Feasibility} && (1.4a)\\ A^T y + z &= c, && \textit{Dual Feasibility} && (1.4b)\\ x_i z_i &= 0, && \textit{Complementarity} && (1.4c)\\ (x, z) &\geqslant 0. && \textit{Non-negativity} && (1.4d) \end{aligned}$$

Theorem 1.3. ***Dual Optimality Conditions*** *The point $(y, z) \in \mathbb{R}^{m+n}$ is a solution of (D) if only if there exist a point $x \in \mathbb{R}^n$ such that the conditions (1.4) hold for (x, y, z).*

In other words, Theorem 1.2 and Theorem 1.3 say the point (x^*, y^*, z^*) solves the system (1.4) if and only if x^* solves the primal problem (P) and (y^*, z^*) solves the dual problem (D). Thus, the point (x^*, y^*, z^*) is called a *primal-dual solution*. We say that (x, y, z) is an primal-dual interior point if $(x, z) > 0$.

A most useful and well-known methods to solve LP problems are **Simplex method** and **Interior-point method**. Simplex method, roughly speaking, consists in to search the solution through the vertices of the feasible region until attain the desirable point that minimizes the objective function. On the other hand, interior-point method searches for a solution in the interior of the feasible region, following a direction that optimizes the objective function. However, since the solution approaches from the interior it never lies on the boundary.

Several analyses and numerical tests show that, for real world and large scale problems, interior-point method is often more efficient and it converges to a solution in few iterations. Thus, we will present a brief review about this method.

1.2 Interior-point Methods

During the last decades the interior-point method has been one of the most used solution methods to solve successfully mathematical programming problems. This method belongs to the iterative class of methods and is characterized by the search for the optimal solution in the interior of the feasible region. This search can be made iterating on the

primal variables, on the dual variables or on the primal-dual variables. The Primal-Dual Methods are known as the most efficient practical approaches [28], therefore we will focus on them along this section.

Consider the primal problem (1.2) and the dual problem (1.3). Define the function $F : \mathbb{R}^{2n+m} \to \mathbb{R}^{2n+m}$ as

$$F(x, y, z) = \begin{pmatrix} Ax - b \\ A^T y + z - c \\ XZe \end{pmatrix},$$

where $X = diag(x_1, \ldots, x_n)$, $Z = diag(z_1, \ldots, z_n)$ and e the vector of all ones. From the optimality conditions (1.4),

$$F(x, y, z) = 0 \text{ with } (x, z) \geqslant 0. \tag{1.5}$$

Primal-dual interior-point method finds solutions (x, y, z) of the non-linear system (1.5) applying a variant of the Newton method [3] generating on each iteration a search direction and determining step lengths in order to guarantee that $(x, z) > 0$, that is, the current solution (x, y, z) be an interior-point.

On the k^{th}-iteration, let (x^k, y^k, z^k) be an interior-point, the Newton search direction $(dx\ dy\ dz)^T$ is obtained solving the following linear system

$$J(x, y, z) \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = -F(x, y, z),$$

where $J(x, y, z)$ is the Jacobian of F:

$$J(x, y, z) = \begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix}.$$

Considering that, for most problems, finding an initial feasible point is a difficult task, the Newton system can be rewritten as

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} r_P \\ r_D \\ -XZe \end{pmatrix},$$

where

$$r_P = -(Ax^k - b), \tag{1.6a}$$

$$r_D = -(A^T y^k + z^k - c), \tag{1.6b}$$

represent the residues of the conditions (1.4a) and (1.4b).

Taking a full step along this direction could generate a non-interior-point, thus the next iteration is computed as

$$(x, y, z) + \alpha(dx, dy, dz), \text{ for some } \alpha \in (0, 1].$$

Now, solving the Newton system we obtain the next set of equations

$$\begin{aligned} (AZ^{-1}XA^T)dy &= r_P + AZ^{-1}Xr_D - AZ^{-1}r_a, && \text{(1.7a)} \\ dx &= Z^{-1}X(A^Tdy - r_D + X^{-1}r_a), && \text{(1.7b)} \\ dz &= X^{-1}(r_a - Zdx), && \text{(1.7c)} \end{aligned}$$

where $r_a = -XZe$ and the existence of X^{-1} and Z^{-1} are guaranteed since $(x, z) > 0$.

In order to maintain the non-negativity of the variables x and z, we take a reduced step if necessary. These step lengths are computed as:

$$\alpha_p = min\left\{1; \tau min\{-\frac{x_j}{dx_j} : \ dx_j < 0\}\right\}, \quad \tau \in (0, 1), \tag{1.8}$$

$$\alpha_d = min\left\{1; \tau min\{-\frac{z_j}{dz_j} : \ dz_j < 0\}\right\}, \quad \tau \in (0, 1). \tag{1.9}$$

The Algorithm 1, called Affine Scaling Interior-Point algorithm, shows a summary of the method.

Observe that, α could take too small values to avoid violating of condition $(x, z) > 0$, causing the method presents slow progress. Thus, a variation of the method changes the function F, so that the new point is more centralized, introducing the parameter $\mu > 0$ whose purpose is to keep the components (x, z) from moving to close to the boundary of the non-negative orthant, arising the concept of *Central Path* [28].

The central path is defined as the solution of

$$F_\mu(x, y, z) = \begin{pmatrix} Ax - b \\ A^Ty + z - c \\ XZe - \mu e \end{pmatrix}, \text{ for each } \mu > 0. \tag{1.10}$$

When $\mu \longrightarrow 0$, the central path converges to a primal-dual solution [28]. The computing of the Newton direction is made solving the linear system

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} r_P \\ r_D \\ -XZe + \mu e \end{pmatrix}.$$

Algorithm 1 – Affine Scaling Interior-Point Method.

Data: y^0 and $(x^0, z^0) > 0$, max: maximal iterations
begin
  for $k = 0, 1, \ldots, max$ **do**
    Passo 1: Compute the residues;
    $r_P = b - Ax$;
    $r_D = c - A^T y - z$;
    $r_a = -XZe$;
    Passo 2: Find the search directions;
    Compute dy in $(AZ^{-1}XA^T)dy = r_P + AZ^{-1}Xr_D - AZ^{-1}r_a$;
    Compute dx in $dx = Z^{-1}X(A^T dy - r_D + X^{-1}r_a)$;
    Compute dz in $dz = X^{-1}(r_a - Zdx)$;
    Passo 3: Compute the primal and dual step lengths, α_p and α_d;
    $\alpha_p = min\left\{1; \tau min\{-\frac{x_j}{dx_j} : \ dx_j < 0\}\right\}, \ \ \tau \in (0,1)$;
    $\alpha_d = min\left\{1; \tau min\{-\frac{z_j}{dz_j} : \ dz_j < 0\}\right\}, \ \ \tau \in (0,1)$;
    Passo 4: Compute the next feasible point in;
    $x^{k+1} = x^k + \alpha_p dx$;
    $y^{k+1} = y^k + \alpha_d dy$;
    $z^{k+1} = z^k + \alpha_d dz$;
  end
end

On the other hand, at any iteration the values of the pairwise product $x_i z_i$ may be reduced to zero on the same rate. For that, the new system to be solved is

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} r_P \\ r_D \\ -XZe + \sigma\mu e \end{pmatrix}, \tag{1.11}$$

where $\sigma \in [0, 1]$ is called *centering parameter* (this name comes from the property: when $\sigma = 1$ the system (1.11) is the central path) and μ is called *duality measure* and is defined by

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i z_i = \frac{x^T z}{n}. \tag{1.12}$$

With these ideas, the **Primal-Dual Path-Following method** can be defined [28] and the general framework is presented on the Algorithm 2.

1.2.1 Non-linear objective function

There exist problems which can not be formulated by a linear objective function. In these cases it is possible adapt the previous analysis, and do little changes considering a

Algorithm 2 – Primal-Dual Path-Following Interior-Point Method.

Data: y^0 and $(x^0, z^0) > 0$, max: maximal iterations
begin
  for $k = 0, 1, \ldots, max$ **do**
    Passo 1: Compute the residues;
    $r_P = b - Ax$;
    $r_D = c - A^T y - z$;
    Passo 2: Compute the search directions dx, dy and dz solving the system
    $$\begin{pmatrix} A & 0 & 0 \\ 0 & A^T & I \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} r_P \\ r_D \\ -XZe + \sigma\mu e \end{pmatrix};$$
    where $\sigma \in [0, 1]$ and μ computed from $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i z_i = \frac{x^T z}{n}$;
    Passo 3: Compute the primal and dual step lengths, α_p and α_d;
    $$\alpha_p = min\left\{1; \tau min\{-\frac{x_j}{dx_j} : \ dx_j < 0\}\right\}, \ \ \tau \in (0, 1);$$
    $$\alpha_d = min\left\{1; \tau min\{-\frac{z_j}{dz_j} : \ dz_j < 0\}\right\}, \ \ \tau \in (0, 1);$$
    Passo 4: Compute the next feasible point ;
    $x^{k+1} = x^k + \alpha_p dx$;
    $y^{k+1} = y^k + \alpha_d dy$;
    $z^{k+1} = z^k + \alpha_d dz$;
  end
end

non-linear objective function $f : \mathbb{R}^n \to \mathbb{R}$. The problem with non-linear objective function and linear constraints is

$$(NLP) \begin{cases} \min & f(x) \\ s.t & Ax = b, \\ & x \geqslant 0. \end{cases}$$

The constraints $x \geqslant 0$ could be replaced by the term $\sum_{i=1}^{n} \ln(x_i)$, called **logarithmic barrier**, because it prevents x_i from being negative. Thus, introducing this term on the formulation (NLP) with a parameter $\mu > 0$, we obtain

$$\begin{cases} \min & f(x) - \mu \sum_{i=1}^{n} \ln(x_i) \\ s.t & Ax = b. \end{cases}$$

Now, let us define the Lagrangean function of the above problem

$$\mathcal{L}(x; y) = f(x) - \mu \sum_{i=1}^{n} \ln(x_i) - y^T(Ax - b), \tag{1.13}$$

where $y \in \mathbb{R}^m$ is the Lagrangean multiplier. The use of the same variable y and parameter μ is purposely named to notice the relation between the Primal-Dual methods and the Logarithmic-barrier methods. Therefore, solving the (NLP) problem is equivalent tor solving for each $\mu > 0$, the minimizing problem

$$\min \qquad \mathcal{L}(x; y).$$

Firstly, we have to find the stationary points of $\mathcal{L}$, that is, (x, y) such that $\nabla\mathcal{L}(x; y) = 0$, then

$$\nabla_x\mathcal{L}(x; y) = \nabla f(x) - \mu X^{-1}e - A^T y = 0, \tag{1.14a}$$

$$\nabla_y\mathcal{L}(x; y) = -(Ax - b) = 0. \tag{1.14b}$$

Let $z \geqslant 0$ be the complementary variable such that $z = \mu X^{-1}e$, then $XZe = \mu e$. Thus, the system (1.14) with the additional condition $XZe - \mu e = 0$, gives the KKT-conditions for the Lagrangean problem:

$$\begin{aligned} \nabla f(x) - z - A^T y &= 0, \\ Ax - b &= 0, \\ XZe - \mu e &= 0. \end{aligned}$$

Defining $F_\mu : \mathbb{R}^{2n+2m} \to \mathbb{R}^{2n+2m}$ as

$$F_\mu(x; y, z) = \begin{pmatrix} \nabla f(x) - z - A^T y \\ Ax - b \\ XZe - \mu e \end{pmatrix}. \tag{1.15}$$

The search direction d is obtained applying the Newton method on the non-linear system $F_\mu(x; y, z) = 0$.

The Jacobian of F_μ is given by

$$J_\mu(x; y, z) = \begin{pmatrix} H & -A^T & -I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix},$$

where $H \in \mathbb{R}^{n\times n}$ denotes the Hessian matrix of f. Then we have

$$\begin{pmatrix} H & -A^T & -I \\ A & 0 & 0 \\ Z & 0 & X \end{pmatrix} \begin{pmatrix} dx \\ dy \\ dz \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix},$$

where

$$r_1 = A^T y^k + z^k - \nabla f(x), \tag{1.16a}$$

$$r_2 = b - Ax^k, \tag{1.16b}$$

$$r_3 = \mu e - xze. \tag{1.16c}$$

Eliminating $dz = X^{-1}(r_3 - Zdx)$ we obtain

$$\begin{pmatrix} H + X^{-1}Z & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} r_1 + X^{-1}r_3 \\ r_2 \end{pmatrix}, \tag{1.17}$$

called the *augmented system*. Generally, as H represents the Hessian and is a symmetric matrix, the system is solved using iterative methods or direct methods, depending on the additional characteristics of H.

We present, in Algorithm 3 the summary for the Primal-dual interior-point method for (NLP) formulation. It is important to point out that the Hessian matrix must be computed at each iteration for solving the system 1.17 and leads to higher computational times.

Algorithm 3 – Primal-Dual Interior-Point Method with non-linear objective function

Data: y^0 and $(x^0, z^0) > 0$
begin
  for $k = 0, 1, \ldots, max$ **do**
    Passo 1: Compute the residues;
    $r_1 = A^T y + z - \nabla f(x)$;
    $r_2 = b - Ax$;
    $r_3 = \mu e - XZe$;
    Passo 2: Compute de search directions;
    $dz = X^{-1}(r_3 - Zdx)$;
    dx and dy solving the system;
    $\begin{pmatrix} H + X^{-1}Z & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} r_1 + X^{-1}r_3 \\ r_2 \end{pmatrix}$
    with $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i z_i = \frac{x^T z}{n}$;
    Passo 3: Compute the primal and dual step lengths, α_p and α_d;
    $\alpha_p = min\left\{1; \tau min\{-\frac{x_j}{dx_j} : dx_j < 0\}\right\}, \quad \tau \in (0,1)$;
    $\alpha_d = min\left\{1; \tau min\{-\frac{z_j}{dz_j} : dz_j < 0\}\right\}, \quad \tau \in (0,1)$;
    Passo 4: Compute the next feasible point;
    $x^{k+1} = x^k + \alpha_p dx$;
    $y^{k+1} = y^k + \alpha_d dy$;
    $z^{k+1} = z^k + \alpha_d dz$;
  end
end

2 Fuzzy Theory

Fuzzy theory, that has been proposed by Lofti A. Zadeh in 1965 [29] considers human subjectivity and imprecision based on the intuitive reasoning, described in the natural language.

This chapter introduces the basic definitions and properties concerning the fuzzy sets necessary to understand the surprise function method, and, the fuzzy optimization formulation.

2.1 Fuzzy Sets

A *crisp set* is defined by a function which only accepts the values 1, meaning that an element fully belongs to a set, and 0, when it does not belongs to a set. Since values between 0 and 1 are also accepted, a fuzzy set is an extension of crisp sets and is defined, by Zadeh [29], as:

Definition 2.1. *Let Ω be a given universal set. A fuzzy set $\mathcal{F}$ on Ω is characterized by a membership function, $\mu_{\mathcal{F}}(x)$, which associates with each point in Ω a real number in $[0,1]$. The value of $\mu_{\mathcal{F}}(x)$ at x represents the membership degree (or grade of membership) of an element of Ω in $\mathcal{F}$. Symbolically,*

$$\mu_{\mathcal{F}} : \Omega \longrightarrow [0,1].$$

2.1.1 Some properties of fuzzy sets

We introduce definitions that are important to characterize another useful concept of this work [27].

Definition 2.2. *A fuzzy set $\mathcal{F}$, defined on Ω, is said to be **empty** if*

$$\mathcal{F} = \varnothing \Longleftrightarrow \mu_{\mathcal{F}}(x) = 0, \ \forall x \in \Omega.$$

Definition 2.3. *The **support** of a fuzzy set $\mathcal{F}$, defined on Ω, is a crisp subset of Ω defined by*

$$supp(\mathcal{F}) = \{x \in \Omega : \mu_{\mathcal{F}}(x) > 0\}.$$

Definition 2.4. *The **height** of a fuzzy set $\mathcal{F}$, defined on a bounded set Ω, is defined by*

$$Hgt(\mathcal{F}) = \max_{x \in \Omega} \{\mu_{\mathcal{F}}(x)\}.$$

That is, it is the maximum membership degree of all the elements of Ω.

Definition 2.5. *A fuzzy set $\mathcal{F}$, defined on Ω, is **normalized** if only if*

$$\exists x \in \mathcal{F}, \;\; \mu_{\mathcal{F}}(x) = Hgt(\mathcal{F}) = 1.$$

*That is, $\mathcal{F}$ is said normalized (or **normal fuzzy set**) if at least one element of Ω has a membership degree 1.*

Definition 2.6. *The α**-cut** of a fuzzy set $\mathcal{F}$, defined on Ω, is a crisp subset of Ω that satisfies*

$$\mathcal{F}_\alpha = \{x \in \Omega : \mu_{\mathcal{F}}(x) \geqslant \alpha, \; \alpha \in [0,1]\}.$$

Definition 2.7. *A fuzzy set $\mathcal{F}$, defined on Ω, is **convex** if*

$$\forall x, y \in \Omega, \forall \lambda \in [0,1] : \mu_{\mathcal{F}}(\lambda x + (1-\lambda)y) \geqslant \min\{\mu_{\mathcal{F}}(x), \mu_{\mathcal{F}}(y)\}.$$

2.1.2 Fuzzy Numbers

A special type of fuzzy sets are those that are defined on $\mathbb{R}$, that is, the membership function is

$$\mu : \mathbb{R} \longrightarrow [0,1].$$

Under certain conditions, as will be seen in the following definition, this function can be viewed as fuzzy number or fuzzy interval [27].

Definition 2.8. *A **fuzzy number** $\mathcal{F}$ is a fuzzy subset of the real numbers $\mathbb{R}$ if satisfies:*

- *$\mathcal{F}$ is a normal set,*
- *$\mathcal{F}_\alpha$ is a closed interval for every $\alpha \in (0,1]$,*
- *The support of $\mathcal{F}$, $supp(\mathcal{F})$, is bounded.*

Some classical fuzzy numbers are defined by particular membership functions:

Example 2.1. *Triangular fuzzy number $\mathcal{F} = (a,b,c)$*

$$\mu(x) = \begin{cases} \dfrac{x-a}{b-a} & , \quad \text{if } x \in [a,b], \\ 1 & , \quad \text{if } x = b, \\ \dfrac{c-x}{c-b} & , \quad \text{if } x \in [b,c], \\ 0 & , \quad \text{otherwise.} \end{cases} \tag{2.1}$$

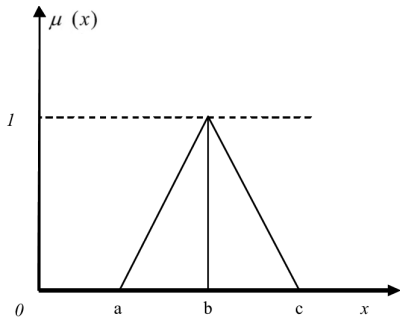


Figure 1 – Representation of triangular fuzzy number.

Example 2.2. *Trapezoidal fuzzy number* $\mathcal{F} = (a, b, c, d)$

$$\mu(x) = \begin{cases} \dfrac{x-a}{b-a} & , \quad \textit{if } x \in [a,b], \\ 1 & , \quad \textit{if } x \in [b,c], \\ \dfrac{d-x}{d-c} & , \quad \textit{if } x \in [c,d], \\ 0 & , \quad \textit{otherwise.} \end{cases} \tag{2.2}$$

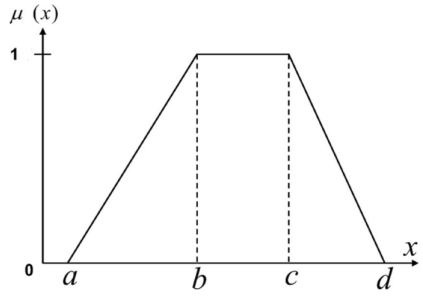


Figure 2 – Representation of trapezoidal fuzzy number.

2.2 Fuzzy sets in terms of Surprise functions

Fuzzy set theory consists in catch imprecise information of the real world and transform it in a numerical value between 0 and 1. If a statement E is true for a particular value x then there is no surprise

$$\mathcal{S}(x|E) = 0, \text{ if } E(x) \text{ is true.}$$

If E is false, then the occurrence of x, being impossible, would be infinitely surprising,

$$\mathcal{S}(x|E) = \infty, \text{ if } E(x) \text{ is false.}$$

Thus surprise is understood as an amount of falseness of the statement E after knowing x.

To connect the fuzzy set approach with the surprise, Neumaier [20] quantifies the evidence for a value x given E by a **degree of consistency** $d(x|E) \in [0,1]$, with $d(x|E) = 0$ if x is impossible given E, and $d(x|E) = 1$ if x is fully consistent with E. It translates any surprise value $\mathcal{S}$ into a degree of consistency d by the strictly monotone decreasing transformation, from $[0,\infty]$ onto $[0,1]$,

$$d = \frac{1}{\mathcal{S}^e + 1}, \text{ for some } e > 0. \tag{2.3}$$

Conversely, from (2.3),

$$\mathcal{S} = \left(\frac{1}{d} - 1\right)^{1/e}, \tag{2.4}$$

where the most useful and recommended exponent is $e = 1/2$ [20].

Therefore, since each statement E is a fuzzy set with membership function $\mu_E(x) = d(x|E)$, the surprise function on x given E is

$$\mathcal{S}_E(x) = \left(\frac{1}{\mu_E(x)} - 1\right)^2. \tag{2.5}$$

2.3 Fuzzy Optimization Problems

In this section, we introduce a type of fuzzy representation for the classical formulation of optimization linear problems (1.1).

Optimization under uncertainty is used to mean optimization when at least one element of the input data, in formulation (1.1), is a real-valued interval, a real-valued random variable, a real-valued fuzzy number, or a real-valued number described by a possibility distribution. The fuzziness can be considered in different ways, for example: the matrix of constraints A, right-hand vector b or vector of costs c as being fuzzy numbers, "$\leqslant$" as fuzzy sets. Another type of fuzziness, in which more general cases are considered, can be seen in [16], [23], [24], [31], [32].

This work focuses on the special case when b is a fuzzy number. It means that the set of constraints belongs to a "real-valued interval".

A mathematical representation correspondent to fuzzy linear optimization is:

$$\begin{aligned} \min \quad & c^T x \\ s.t \quad & Ax \leqslant \tilde{b}, \\ & x \geqslant 0, \end{aligned} \tag{2.6}$$

where $\tilde{b}$ is a vector in $\mathbb{R}^m$ whose components are fuzzy numbers, $A \in \mathbb{R}^{m\times n}$ is the matrix of constraints. Fuzzy numbers are fuzzy sets, each component of Ax ($\in \mathbb{R}^m$) has a degree of belonging to a subset of $\mathbb{R}$. This type of formulation, in which the right-hand side is a fuzzy number will be treated below using the theory of surprise functions [18], [20].

2.3.1 Fuzzy optimization using surprise functions

The surprise function approach searches for the best solution in the constraints and maximizes it on all combined α-levels applying a dynamical penalty to the violations of constraints by surprise functions as follows [20]:

Let $A = [a_{ij}]_{i=1,\dots,m}^{j=1,\dots,n}$ be, for each $i = 1, \dots, m$

$$\sum_{j=1}^{n} a_{ij}x_j \leqslant \tilde{b_i} \Longleftrightarrow \sum_{j=1}^{n} a_{ij}x_j = \tilde{\xi_i}, \tag{2.7}$$

where

$$\mu_i(\xi_i) = pos(\tilde{b_i} \geqslant \tilde{\xi_i}).$$

and the membership function $\mu_i()$ is defined using the possibility distribution $pos()$.

Once the membership function known, by (2.5), we are able to formulate the surprise function optimization:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} \mathcal{S}_i \left(\sum_{j=1}^{n} a_{ij}x_j \right) \\ s.t. \quad & 0 \leqslant x \leqslant U. \end{aligned} \tag{2.8}$$

with $U \in \mathbb{R}^n$ is the upper limit value for variable x. The objective function minimizes the sum of the surprise function evaluated on each restriction. The single constraint provides the domain of the variable. We will adapt this formulation, for the particular case in which $\tilde{b}_i$ is triangular fuzzy number (see Example 2.1), for solving the dose distribution problem in radiation therapy planning. And since μ represents the membership function associated with the fuzzy-number $\tilde{b}$, we have (see subsection 2.1.1) for all $x, y \in \mathbb{R}$ and for all $\lambda \in [0, 1]$

$$\mu(\lambda x + (1-\lambda)y) \geqslant \min\{\mu(x), \mu(y)\},$$

thus

$$\mu(x) \leqslant \mu(\lambda x + (1-\lambda)y) \quad \Rightarrow \quad \left(\frac{1}{\mu(x)} - 1\right)^2 \geqslant \left(\frac{1}{\mu(\lambda x + (1-\lambda)y)} - 1\right)^2 \tag{2.9}$$

$$\mu(y) \leqslant \mu_i(\lambda x + (1-\lambda)y) \quad \Rightarrow \quad \left(\frac{1}{\mu_i(y)} - 1\right)^2 \geqslant \left(\frac{1}{\mu(\lambda x + (1-\lambda)y)} - 1\right)^2 \tag{2.10}$$

multiplying (2.9) by λ and (2.10) by $(1-\lambda)$, and then summing both inequalities, we get

$$\lambda\mathcal{S}(x) + (1-\lambda)\mathcal{S}(y) \geqslant \mathcal{S}(\lambda x + (1-\lambda)y).$$

Therefore, the surprise function defined from the membership function of a fuzzy number is a convex function.

3 Radiation Therapy

The objective of radiation therapy is to deliver a dose of radiation in order to eliminate tumor cells avoiding, as much as possible, the nearby healthy organs and tissues.

3.1 General Description

Radiation therapy is a technique used for treating cancer patients with ionic radiation. We consider a machine called ***linear accelerator*** (see Figure 3[1]) for the external radiation therapy method. The goal of this machine consists in delivering radiation from the **collimator** in different forms and intensities, this radiation is shaped by beams, whose discretisation is called *beamlets*.

Figure 3 – Linear Accelerator.

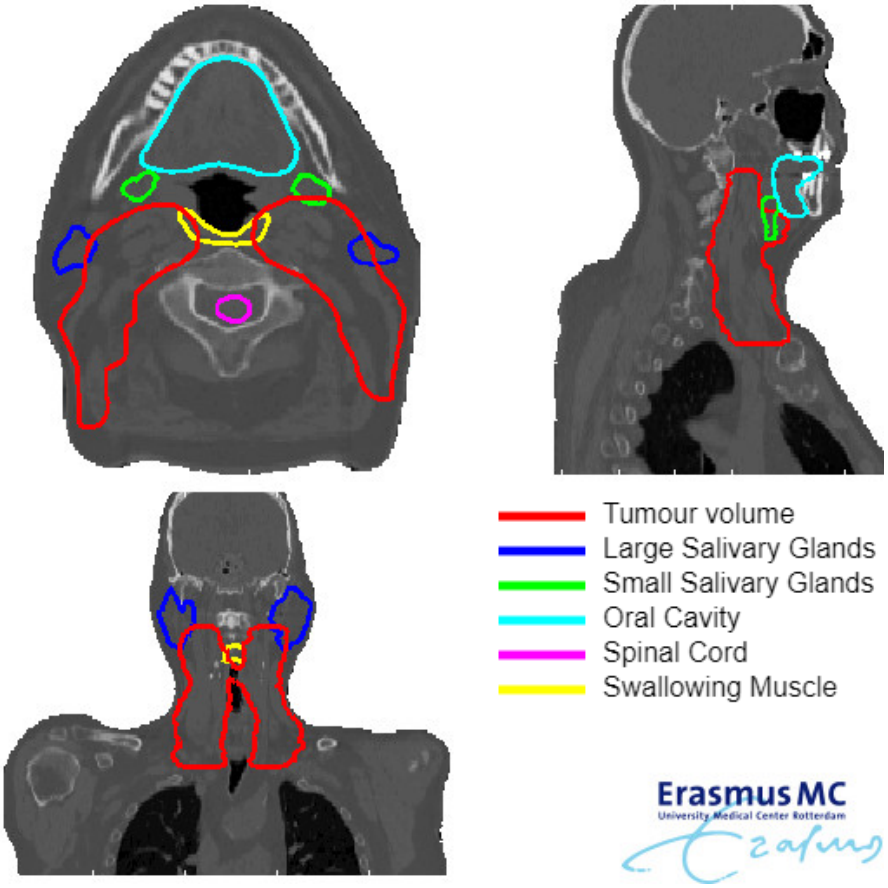
As the head of the linear accelerator can be rotated around the patient, the treatment provides radiation in a wide range of directions. The amount of radiation absorbed by the tissue is called **dose** and its unit is **Gray** (**Gy**), where $1Gy = 1J/Kg$.

Therefore, is important to create a plan of distribution of radiation such that cancer cells receive the highest radiation concentration so that they can be eliminated, taking into account that healthy tissue as well as the organs, which can be close to the

[1] Font: *RADIOLOGY ONCOLOGY SYSTEMS*. https://www.oncologysystems.com/medical-equipment/radiation-therapy. Access date: 24-08-2020.

tumor or in the direction of the beams, should have no severe damage or induce future complications.

Figure 4 – Example of CT-scan.



A good treatment plan consist of, firstly, images of the patient focusing on the tumor and the surrounding tissues. To achieve this, a CT-scan (Computed Tomography) is made to delineate and locate the tumor, the organs-at-risk and other structures (see Figure 4[2]). Then, these images are discretized and each element is called *pixel*.

Note that, when a set of CT-scans (of the same region) is considered, a volumetric representation of the patient is obtained and the pixels take on a 3D-interpretation: *voxels*. With the localization of the volume of tumor just known, the radiation oncologist determines how much radiation deliver to the tumor and the tolerance values for the other tissues. This planning will be performed with the inverse planning, whose essence is an optimization model that find the best solution satisfying the minimal requirements for tumor and tolerance for the other tissues.

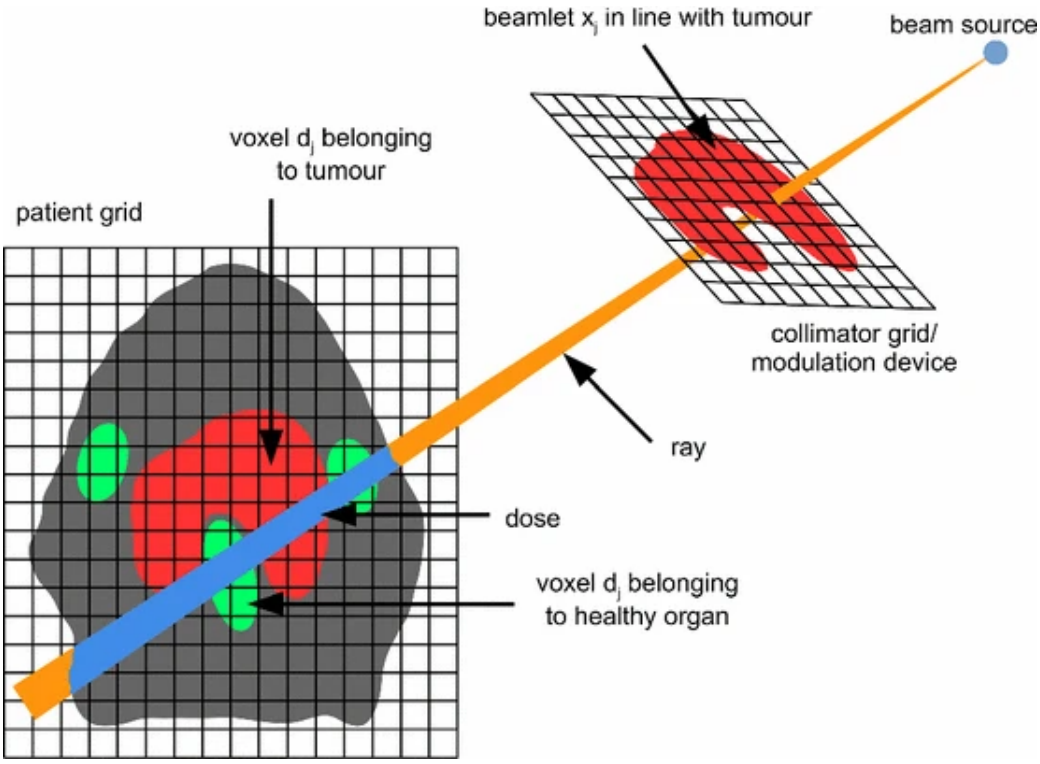
[2] Font: Sebastiaan Breedveld assistant professor at the unit Medical Physics, department of Radiation Oncology Erasmus MC Cancer Institute (Erasmus University Medical Center) Rotterdam, The Netherlands. https://sebastiaanbreedveld.nl/radiotherapy.html.

3.2 Radiotherapy Optimization

Radiation therapy planning is designed to we take advantage of mathematical tools such as optimization modeling to formulate the goals of the treatment, in which the problem is to find the best dose distribution. That is, given the number of beams and direction of each beam, we have to find the intensity that is optimal for each beam.

Figure 5 expands Breedveld S. et al. [5] decomposition of the radiotherapy problem:

Figure 5 – Picturing of the radiotherapy decomposition



Ionising radiation originates from the beam source point and falls onto a collimator. The collimation device allows shaping the beam in different forms and intensities, and is discretised in *beamlets*. The longer a beamlet is "open", the higher the intensity through that beamlet, and the higher the resulting dose in the patient. As soon as the beam enters the patient, the ionising radiation interact with the tissue, leading to dose (cell damage). The patient is discretised in *voxels*.

The value of the radiation dose received by the voxel j from the beamlet i is stored in a matrix A called *pencil-beam matrix* or *attenuation matrix*. This matrix relates the dose distribution and the beamlet intensity. In order to optimize the radiation, the decision variable x is represented the beamlet intensity.

As each region of the patient needs different doses, the matrix A and the dose vector are separated according to the region of interest.

One linear programming formulation is to minimize the total integral dose subject to a lower bound on the dose to the tumor. The integral dose is the total dose summed over all of the pixels.

Thus, a typical formulation is Censor et al. (1988) [9] and Cormack and Quinto (1990) [10]:

$$\begin{array}{ll} \min & f(x) = c^T x \\ s.t & Bx \leqslant b_{body}, \\ & C_i x \leqslant c_i, \ \forall i = 1, \ldots, N, \\ & Tx \geqslant t_{min}, \\ & Tx \leqslant t_{max}, \\ & 0 \leqslant x \leqslant U. \end{array} \tag{3.1}$$

where

- the rows of B represent the body voxels, C_i represent the critical organs voxels and T represent the tumor voxels,
- the vectors b_{body} and c_i for $i = 1, \ldots, N$ are the maximum value allowed for the dose of the healthy tissue and the critical organs, respectively,
- t_{min} is the minimum dose expected for cause damage on the tumor cells and t_{max} is the maximum dose that avoiding several consequences on the surrounding tissues,
- since it is no possible to deliver negative radiation, $x \geqslant 0$ and it is limited by U.

Taking

$$b = \begin{pmatrix} b_{body} \\ c_1 \\ \vdots \\ c_N \\ -t_{min} \\ t_{max} \end{pmatrix} \quad and \quad A = \begin{pmatrix} B \\ C_1 \\ \vdots \\ C_N \\ -T \\ T \end{pmatrix}.$$

A simplified linear programming formulation is

$$\begin{array}{ll} \min & c^T x \\ s.t & Ax \leqslant b, \\ & 0 \leqslant x \leqslant U. \end{array} \tag{3.2}$$

where $A \in \mathbb{R}^{m \times n}$ is the pencil-beam matrix such that if the beamlet i intersects a voxel j, a_{ij} is the positive fraction of the area of the intersection, that is $a_{ij} \in [0, 1]$, for all $i = 1, \ldots, m$, $j = 1, \ldots, n$.

4 Problem Formulation

The problem we will solve in this work is the fuzzy formulation (2.8) of the radiotherapy problem optimization (3.2). That is, we are going to minimize the surprise value of the radiation dose delivered on each cell of the patient.

$$\begin{aligned} \min \quad & \sum_{i=1}^{TP} \mathcal{S}_i \left(\sum_{j=1}^{N} a_{ij} x_j \right) \\ s.t. \quad & 0 \leqslant x \leqslant U, \end{aligned}$$

where $A := [a_{ij}]_{i=1,\dots TP,\ j=1,\dots N}$ represents the pencil-beam matrix, TP is the total of voxels and N the total of beamlets and

$$\mathcal{S}_i(\xi_i) = \left(\frac{1}{\mu_i(\xi_i)} - 1 \right)^2, \ with\ \xi_i = \sum_{j=1}^{N} a_{ij} x_j. \tag{4.1}$$

and, for each $i = 1, \cdots, TP$, $\mathcal{S}_i$ is a convex function, then for all $x, y \in \mathbb{R}$, and $\forall \lambda \in [0, 1]$

$$\begin{aligned} \sum_{i=1}^{TP} \mathcal{S}_i(\lambda x + (1-\lambda) y) &\leqslant \sum_{i=1}^{TP} [\lambda \mathcal{S}_i(x) + (1-\lambda)\mathcal{S}_i(y)] \\ &= \sum_{i=1}^{TP} \lambda \mathcal{S}_i(x) + \sum_{i=1}^{TP} (1-\lambda)\mathcal{S}_i(y) \\ &= \lambda \sum_{i=1}^{TP} \mathcal{S}_i(x) + (1-\lambda) \sum_{i=1}^{TP} \mathcal{S}_i(y). \end{aligned}$$

$$\Rightarrow \sum_{i=1}^{TP} \mathcal{S}_i(\lambda x + (1-\lambda) y) \leqslant \lambda \sum_{i=1}^{TP} \mathcal{S}_i(x) + (1-\lambda) \sum_{i=1}^{TP} \mathcal{S}_i(y).$$

Therefore, we obtain a convex function.

This problem has linear constraints and nonlinear convex objective function [20], therefore we will apply the method described in subsection 1.2.1.

Throughout this chapter we compute the gradient vector and Hessian matrix of the objective function defined by the surprise function and, considering b in (3.2) as a vector of triangular fuzzy numbers, we use its respective membership function to express the gradient and Hessian matrix on a matricial form. Then we add constraints in order to ensure a good quality of solutions and a good performance for the developed algorithm based on Algorithm 3.

4.1 Computing the gradient and Hessian Matrix of the objective function

The objective function is defined as:

$$f(x) = \sum_{i=1}^{TP} \mathcal{S}_i \left(\sum_{j=1}^{N} a_{ij} x_j \right), \tag{4.2}$$

First, we compute the partial derivates of f. For each $k = 1, \ldots, N$,

$$\frac{\partial f(x)}{\partial x_k} = \sum_{i=1}^{TP} \frac{\partial \mathcal{S}_i(\xi_i)}{\partial x_k}.$$

For each $i = 1, \ldots, TP$ and using the rule chain, we have

$$\frac{\partial \mathcal{S}_i(\xi_i)}{\partial x_k} = \frac{\partial \mathcal{S}_i(\xi_i)}{\partial \xi_i} \frac{\partial \xi_i}{\partial x_k}, \tag{4.3}$$

where

$$\begin{aligned} \frac{\partial \mathcal{S}_i(\xi_i)}{\partial \xi_i} &= \frac{\partial}{\partial \xi_i} \left(\frac{1}{\mu_i(\xi_i)} - 1 \right)^2 \\ &= 2 \left(\frac{1}{\mu_i(\xi_i)} - 1 \right) \left(-\mu_i(\xi_i)^{-2} \frac{\partial \mu_i(\xi_i)}{\partial \xi_i} \right) \\ &= -2 \frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} \frac{\partial \mu_i(\xi_i)}{\partial \xi_i}. \end{aligned}$$

Replacing it in (4.3) we obtain

$$\begin{aligned} \frac{\partial \mathcal{S}_i(\xi_i)}{\partial x_k} &= \left[-2 \frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} \frac{\partial \mu_i(\xi_i)}{\partial \xi_i} \right] \frac{\partial \xi_i}{\partial x_k} \\ &= -2 \frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} \frac{\partial \mu_i(\xi_i)}{\partial x_k}. \end{aligned} \tag{4.4}$$

Then

$$\frac{\partial f(x)}{\partial x_k} = -2 \sum_{i=1}^{TP} \frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} \frac{\partial \mu_i(\xi_i)}{\partial x_k}. \tag{4.5}$$

Now, we compute the second-order partial derivate of f. For each $p = 1, \ldots, N$,

$$\frac{\partial^2 f(x)}{\partial x_p \partial x_k} = \sum_{i=1}^{TP} \frac{\partial}{\partial x_p} \left(\frac{\partial \mathcal{S}_i(\xi i)}{\partial x_k} \right).$$

From (4.4) and, in order to simplify the calculations, using the following equality

$$\frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} = \mu_i(\xi_i)^{-3} - \mu_i(\xi_i)^{-2},$$

we obtain

$$\begin{aligned}\frac{\partial^2 f(x)}{\partial x_p \partial x_k} &= -2\sum_{i=1}^{TP}\frac{\partial}{\partial x_p}\left[\frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2}\frac{\partial \mu_i(\xi_i)}{\partial x_k}\right] \\ &= -2\sum_{i=1}^{TP}\left[\frac{\partial}{\partial x_p}\left(\mu_i(\xi_i)^{-3}-\mu_i(\xi_i)^{-2}\right)\frac{\partial \mu_i(\xi_i)}{\partial x_k}+\frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2}\frac{\partial^2 \mu_i(\xi_i)}{\partial x_p \partial x_k}\right] \\ &= 2\sum_{i=1}^{TP}\left[M_i(\xi_i)\frac{\partial \mu_i(\xi_i)}{\partial x_p}\frac{\partial \mu_i(\xi_i)}{\partial x_k}-\frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2}\frac{\partial^2 \mu_i(\xi_i)}{\partial x_p \partial x_k}\right], \end{aligned} \tag{4.6}$$

where

$$M_i(\xi_i) = 3\mu_i(\xi_i)^{-4} - 2\mu_i(\xi_i)^{-3} > 0, \text{ for each } i = 1, \ldots, TP. \tag{4.7}$$

Therefore, the gradient vector and Hessian matrix of f are, respectively,

$$\nabla f(x) = -2\left[\sum_{i=1}^{TP}\frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2}\frac{\partial \mu_i(\xi_i)}{\partial x_k}\right]_{k=1,\ldots,N}, \tag{4.8}$$

and

$$\mathbf{H}(x) = 2\left[\sum_{i=1}^{TP}\left(M_i(\xi_i)\frac{\partial \mu_i(\xi_i)}{\partial x_p}\frac{\partial \mu_i(\xi_i)}{\partial x_k}-\frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2}\frac{\partial^2 \mu_i(\xi_i)}{\partial x_p \partial x_k}\right)\right]_{p,k=1\ldots,N}. \tag{4.9}$$

4.2 Fuzzy Optimization applied to Radiotherapy

The objective function (4.2) is minimize the sum of surprise functions and they assume small values when the membership functions are closed to 1, that is, the variables ξ have a high membership degree in $\tilde{b}$. There exist a great variety of membership functions, μ for the radiotherapy problem (see chapter 1), an attractive option is taking $\tilde{b}$ a trapezoidal fuzzy number since would be exist a larger set of values for ξ that have a high membership degree in $\tilde{b}$ than when $\tilde{b}$ is a triangular fuzzy number (see Example 2.1 and Example 2.2). However, in the compute of the derivates, the set in which the trapezoidal fuzzy number is zero is larger than the triangular fuzzy number, increasing the possibility of maintaining the current feasible solution in same position (see Algorithm 3). Therefore we considering $\tilde{b}$ as a vector of TP-triangular fuzzy numbers, ie, $\tilde{b} = (\tilde{b}_i)_{i=1,\ldots,TP}$ with $\tilde{b}_i = (b_i^1 \;\; b_i \;\; b_i^2)$, the membership function for the i-th constraint of the problem is

$$\mu_i(\xi_i) = \begin{cases} \dfrac{\xi_i - b_i^1}{b_i - b_i^1}, & \text{if } \xi_i \in [b_i^1, b_i), \\ 1\,, & \text{if } \xi_i = b_i, \\ \dfrac{b_i^2 - \xi_i}{b_i^2 - b_i}, & \text{if } \xi_i \in (b_i, b_i^2], \\ 0\,, & \text{otherwise.} \end{cases} \tag{4.10}$$

We will take the partial derivates on the open-intervals

$$\frac{\partial \mu_i(\xi_i)}{\partial x_k} = \begin{cases} \dfrac{a_{ik}}{b_i - b_i^1}, & \text{if } \xi_i \in (b_i^1, b_i), \\ \dfrac{-a_{ik}}{b_i^2 - b_i}, & \text{if } \xi_i \in (b_i, b_i^3). \end{cases} \tag{4.11}$$

With these considerations, (4.8) and (4.9) become

$$\begin{aligned} \nabla f(x) &= -2 \left[\sum_{\xi_i \in (b_i^1, b_i)} \frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} \frac{a_{ik}}{b_i - b_i^1} + \sum_{\xi_i \in (b_i, b_i^2)} \frac{\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2} \frac{-a_{ik}}{b_i^2 - b_i} \right]_{k=1,\ldots,N} \\ &= \left[\sum_{\xi_i \in (b_i^1, b_i)} \frac{-2\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2 (b_i - b_i^1)} a_{ik} + \sum_{\xi_i \in (b_i, b_i^2)} \frac{2\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2 (b_i^2 - b_i)} a_{ik} \right]_{k=1,\ldots,N}, \end{aligned} \tag{4.12}$$

and, as $\dfrac{\partial^2 \mu_i(\xi_i)}{\partial x_p \partial x_k} = 0$, for all $p, k = 1, \ldots, N$,

$$\begin{aligned} \mathbf{H}(x) &= 2 \left[\sum_{\xi_i \in (b_i^1, b_i)} M_i(\xi_i) \frac{a_{ik} a_{ip}}{(b_i - b_i^1)^2} + \sum_{\xi_i \in (b_i, b_i^2)} M_i(\xi_i) \frac{a_{ik} a_{ip}}{(b_i^2 - b_i)^2} \right]_{p,k=1\ldots,N} \\ &= \left[\sum_{\xi_i \in (b_i^1, b_i)} a_{ik} \frac{2M_i(\xi_i)}{(b_i - b_i^1)^2} a_{ip} + \sum_{\xi_i \in (b_i, b_i^2)} a_{ik} \frac{2M_i(\xi_i)}{(b_i^2 - b_i)^2} a_{ip} \right]_{p,k=1\ldots,N}. \end{aligned} \tag{4.13}$$

However, (4.12) and (4.13) can be written on a matrix form that helps us to view the relation, of the gradient and Hessian, with A. For that, we define the TP-dimensional vectors $d_{\nabla f}$ and d_H as

$$d_{\nabla f_i} = \begin{cases} \dfrac{-2\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2 (b_i - b_i^1)}, & \text{if } \xi_i \in (b_i^1, b_i), \\ \dfrac{2\mathcal{S}_i(\xi_i)^{1/2}}{\mu_i(\xi_i)^2 (b_i^2 - b_i)}, & \text{if } \xi_i \in (b_i, b_i^2), \\ 0\,, & \text{otherwise.} \end{cases} \tag{4.14}$$

$$d_{H_i} = \begin{cases} \dfrac{2M_i}{(b_i - b_i^1)^2}, & \text{if } \xi_i \in (b_i^1, b_i), \\ \dfrac{2M_i}{(b_i^2 - b_i)^2}, & \text{if } \xi_i \in (b_i, b_i^2), \\ 0\,, & \text{otherwise.} \end{cases} \tag{4.15}$$

for each $i = 1, \ldots, TP$.

Then,

$$\nabla f(x) = A^T d_{\nabla f}, \tag{4.16}$$

$$\mathbf{H}(x) = A^T D_H A, \tag{4.17}$$

where D_H is a $TP \times TP$-matrix such that $D_H = diagonal(d_H)$ and $d_{H_i} \geqslant 0$, from (4.7), then for any $w \in \mathbb{R}^N$

$$w^T H w = w^T A^T D_H A w = (Aw)^T (D_H^{1/2} D_H^{1/2})(Aw) = \|D_H^{1/2}(Aw)\|^2 \geqslant 0,$$

therefore, $\mathbf{H}(x)$ is semi-positive definite matrix.

4.3 Adding constraints and developing the interior-point method

The surprise model (2.8) can be modified in order to ensure a good quality of solutions and a good performance of the algorithm. We will add the following constraints:

- Taking into account relation (2.7), and considering (4.1), ξ will be considered as a variable that must satisfy:
$$Ax = \xi.$$
- Since ξ is a variable, that must satisfy $\xi \leqslant \tilde{b} = (b_1 \;\; b \;\; b_2)$(triangular fuzzy number). Outside of the interval (b_1, b_2) the surprise function takes on large values. On the other hand for
$$b_1 \leqslant \xi \leqslant b_2,$$
the surprise function value be will be smaller and, therefore, $f(x)$ to be minimized.

The previous analysis allows us to define an improved formulation of the problem:

$$\begin{cases} \min & f(x) \\ s.t & x \leqslant U, \\ & Ax = \xi, \\ & \xi \geqslant b_1, \\ & \xi \leqslant b_2, \\ & x \geqslant 0. \end{cases} \tag{4.18}$$

The interior-point method will be developed with this formulation. Adding the slack variables $v, z_1, z_2 \geqslant 0$ and the logarithmic-barrier-terms (4.18) becomes

$$\begin{cases} \min & f(x) \\ s.t & x + v = U, \\ & Ax = \xi, \\ & \xi - z_1 = b_1, \\ & \xi + z_2 = b_2, \\ & x, v, z_1, z_2 \geqslant 0. \end{cases} \iff \begin{cases} \min & \tilde{f}(\overline{x}) \\ s.t & x + v = U, \\ & Ax = \xi, \\ & \xi - z_1 = b_1, \\ & \xi + z_2 = b_2. \end{cases}$$

where

$$\tilde{f}(\overline{x}) = f(x) - \gamma_x \sum_{i=1}^{N} \ln x_i - \gamma_v \sum_{i=1}^{N} \ln v_i - \gamma_{z_1} \sum_{i=1}^{TP} \ln z_i^1 - \gamma_{z_2} \sum_{i=1}^{TP} \ln z_i^2,$$

with $\overline{x} = (x, v, z_1, z_2)^T$.

Consider the Lagrangean function of $\tilde{f}$ as follows

$$\mathcal{L}(\overline{x}; \overline{y}) = \tilde{f}(x) - y^T(x + v - U) - w_1^T(\xi - z_1 - b_1) - w_2^T(\xi + z_2 - b_2) - q^T(Ax - \xi), \quad (4.19)$$

where $\overline{y} = (y, w_1, w_2, q)^T$ is the non-negative dual variable of the formulation (4.18).

Computing the gradient of $\mathcal{L}$ gives raise to:

$$\begin{aligned}
\nabla_x \mathcal{L}(\overline{x}; \overline{y}) &= \nabla f(x) - \gamma_x X^{-1} e - y - A^T q, && (4.20a)\\
\nabla_\xi \mathcal{L}(\overline{x}; \overline{y}) &= -w_1 - w_2 + q, && (4.20b)\\
\nabla_v \mathcal{L}(\overline{x}; \overline{y}) &= -\gamma_v V^{-1} e - y, && (4.20c)\\
\nabla_{z_1} \mathcal{L}(\overline{x}; \overline{y}) &= -\gamma_{z_1} Z_2^{-1} e + w_1, && (4.20d)\\
\nabla_{z_2} \mathcal{L}(\overline{x}; \overline{y}) &= -\gamma_{z_2} Z_2^{-1} e - w_2, && (4.20e)\\
\nabla_y \mathcal{L}(\overline{x}; \overline{y}) &= -(x + v - U), && (4.20f)\\
\nabla_{w_1} \mathcal{L}(\overline{x}; \overline{y}) &= -(\xi - z_1 - b_1), && (4.20g)\\
\nabla_{w_2} \mathcal{L}(\overline{x}; \overline{y}) &= -(\xi + z_2 - b_2), && (4.20h)\\
\nabla_q \mathcal{L}(\overline{x}; \overline{y}) &= -(Ax - \xi), && (4.20i)
\end{aligned}$$

where X, V, Z_1 and Z_2 are diagonal matrix formed by the components of x, v, z_1 and z_2, respectively, and e is the vector of $1's$. Let p, s, t_1 and t_2 be the complementary (non-negative) variables such that

$$\begin{aligned}
p = \gamma_x X^{-1} e &\iff XPe = \gamma_x e, && (4.21a)\\
s = \gamma_v V^{-1} e &\iff VSe = \gamma_v e, && (4.21b)\\
t_1 = \gamma_{z_1} Z_1^{-1} e &\iff Z_1 T_1 e = \gamma_{z_1} e, && (4.21c)\\
t_2 = \gamma_{z_2} Z_2^{-1} e &\iff Z_2 T_2 e = \gamma_{z_2} e. && (4.21d)
\end{aligned}$$

Replacing on the system (4.20), $\nabla\mathcal{L}(\overline{x}; \overline{y}) = 0$ and the additional conditions

(4.21) which are the KKT-conditions of the problem. Thus, let us define the function F

$$F(\overline{x};\overline{y}) = \begin{pmatrix} \nabla f(x) - p - y - A^T q \\ -w_1 - w_2 + q \\ y + s \\ -t_1 + w_1 \\ t_2 + w_2 \\ x + v - U \\ \xi - z_1 - b_1 \\ \xi + z_2 - b_2 \\ Ax - \xi \\ XPe - \gamma_x e \\ VSe - \gamma_v e \\ T_1 Z_1 e - \gamma_{z_1} e \\ T_2 Z_2 e - \gamma_{z_2} e \end{pmatrix}. \tag{4.22}$$

Applying Newton's method to the equation $F(\overline{x};\overline{y}) = 0$ gives

$$\begin{pmatrix} H & 0 & 0 & 0 & 0 & -I & 0 & 0 & -A^T & -I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -I & -I & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & -I & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & I & 0 & 0 & 0 & 0 & I \\ I & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ A & -I & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ P & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & X & 0 & 0 & 0 \\ 0 & 0 & S & 0 & 0 & 0 & 0 & 0 & 0 & 0 & V & 0 & 0 \\ 0 & 0 & 0 & T_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_1 & 0 \\ 0 & 0 & 0 & 0 & T_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & Z_2 \end{pmatrix} \cdot \mathbf{d} = \mathbf{r}, \tag{4.23}$$

where:

- H is the $N \times N$ - Hessian matrix of f, defined in (4.17). It is symmetric and semi-positive definite,

- $\mathbf{d} = \begin{pmatrix} dx \\ d\xi \\ dv \\ dz_1 \\ dz_2 \\ dy \\ dw_1 \\ dw_2 \\ dq \\ dp \\ ds \\ dt_1 \\ dt_2 \end{pmatrix}$ and $r = \begin{pmatrix} p + y + A^T q - \nabla f(x) \\ w_1 + w_2 - q \\ -y - s \\ t_1 - w_1 \\ -t_2 - w_2 \\ U - x - v \\ z_1 + b_1 - \xi \\ b_2 - z_2 - \xi \\ \xi - Ax \\ \gamma_x e - XPe \\ VSe - \gamma_v e \\ \gamma_{z_1} e - T_1 Z_1 e \\ \gamma_{z_2} e - T_2 Z_2 e \end{pmatrix}$.

Solving (4.23) and denoting $X^{-1}P + V^{-1}S$ and $Z_1^{-1}T_1 + Z_2^{-1}T_2$ as D_x and D_z, respectively, we obtain

$$
\begin{aligned}
[H + D_x + A^T D_z A]dx &= r_x, && \text{(4.24a)}\\
d\xi &= Adx - r_9, && \text{(4.24b)}\\
dv &= r_6 - dx, && \text{(4.24c)}\\
dz_1 &= d\xi - r_7, && \text{(4.24d)}\\
dz_2 &= r_8 - d\xi, && \text{(4.24e)}\\
dp &= X^{-1}(r_{10} - Pdx), && \text{(4.24f)}\\
ds &= V^{-1}(r_{11} - Sdx), && \text{(4.24g)}\\
dt_1 &= Z_1^{-1}(r_{12} - T_1 dz_1), && \text{(4.24h)}\\
dt_2 &= Z_2^{-1}(r_{13} - T_2 dz_2), && \text{(4.24i)}\\
dq &= r_2 + dw_1 + dw_2, && \text{(4.24j)}\\
dy &= r_3 - ds, && \text{(4.24k)}\\
dw_1 &= r_4 + dt_1, && \text{(4.24l)}\\
dw_2 &= r_5 - dt_2, && \text{(4.24m)}
\end{aligned}
$$

where

$$
\begin{aligned}
r_x &:= r_1 + A^T(r_2 + r_{w_1} + r_{w_2}) + r_3 + V^{-1}(Sr_6 - r_{11}) + X^{-1}r_{10},\\
r_{w_1} &:= r_4 + Z_1^{-1}[r_{12} + T_1(r_7 + r_9)],\\
r_{w_2} &:= r_5 - Z_2^{-1}[r_{13} - T_2(r_8 + r_9)].
\end{aligned}
$$

The linear system (4.24) allows us to compute the search directions. The main difficult is to find dx since a method for solving a system of N-linear equations must be

used. However, from (4.17), $H = A^T D_H A$, therefore, let $\overline{D} = D_H + D_z$. To compute dx we have to solve the linear system:

$$(A^T \overline{D} A + D_x)dx = r_x. \tag{4.26}$$

When the dataset is obtained from a real data medical, these type of problems are large problems with sparse and full-rank matrices. The matrix $C = A^T \overline{D} A + D_x$ is symmetric, positive definite (because $x > 0$ and $z > 0$) and thus we can solve the system (4.26) applying the Cholesky decomposition. As this system changes at each iteration, we have to solve a new $N \times N$ linear system where the quality of the solution obtained will depend on the matrix A. This step takes a considerable processing time and computational effort, therefore we will use the following strategy

$$C = \hat{C}^T \hat{C} + D_x, \text{ where } \hat{C} = \overline{D}^{1/2} A,$$

and will decompose C using the Cholesky-function of MATLAB.

4.4 Developing the Algorithm

In this section we will present the algorithm implemented. First we explain how determine the step lengths and stopping criterian.

4.4.1 Determining the step lengths

The step length is next computed as (1.8) and (1.9) for the non-negative variables. On the other hand, despite ξ being free, we must to have special control on it when it is close of b (for the triangular fuzzy number $\tilde{b} = (b_1 \;\; b \;\; b_2)$) because the derivate of μ will be zero (4.11). For that, we define

$$\beta = \min\{1, \tau\beta_L, \tau\beta_R\}, \tag{4.27}$$

where β_L and β_R represent the length step when ξ approaches b from the left and from the right side, respectively, and are computed as

$$\beta_L = \min\left\{\frac{b-\xi}{d\xi} : b_1 < \xi < b,\ d\xi > 0\right\} \text{ and } \beta_R = \min\left\{\frac{b-\xi}{d\xi} : b < \xi < b_2,\ d\xi < 0\right\}.$$

Let $\alpha_x, \alpha_{z_1}, \alpha_{z_2}, \alpha_v$ denote the step lengths of x, z_1, z_2, v, respectively, where

$$\alpha_p = \min\{\alpha_x, \alpha_{z_1}, \alpha_{z_2}, \alpha_v, \beta\}, \tag{4.28}$$

is the step length for the primal variables, and $\alpha_{t_1}, \alpha_{t_2}, \alpha_s, \alpha_p$ the step lengths of t_1, t_2, s, p, respectively,

$$\alpha_d = \min\{\alpha_{t_1}, \alpha_{t_2}, \alpha_s, \alpha_p\}, \tag{4.29}$$

is the step length for the dual variables.

4.4.2 Stopping Criterian

Besides having control in not violating the non-negative conditions on the variables, we need that the residues and the *gap* to be reduced to maintain the feasibility of the solution. For that, we will take the relative errors of the main constraints:

$$CP_1 = \frac{\|U - x - v\|}{\|U\| + 1}, \tag{4.30a}$$

$$CP_2 = \frac{\|\xi - Ax\|}{\|x\| + \|\xi\| + 1}, \tag{4.30b}$$

$$CP_3 = \frac{\|b_1 + z_1 - \xi\|}{\|b_1\| + 1}, \tag{4.30c}$$

$$CP_4 = \frac{\|b_2 - z_2 - \xi\|}{\|b_2\| + 1}, \tag{4.30d}$$

$$CP_5 = \frac{\|\nabla f - p - y - A^T q\|}{\|\nabla f\| + 1}, \tag{4.30e}$$

$$CP_6 = \frac{gap}{2f(x) + 1}, \tag{4.30f}$$

where the addition of 1 in the denominators is to avoid division by small values. Thus, we define the stopping criterian as the maximum value:

$$CP = \max\{CP_1,\ CP_2,\ CP_3,\ CP_4,\ CP_5,\ CP_6\}, \tag{4.31}$$

and ask it to satisfy a given tolerance.

The resulting algorithm after the previous analysis is summarized in the Algorithm 4, where are used the following parameters:

$$\begin{aligned} maxit &= 1000, \\ tol &= 10^{-4}, \\ \tau &= 0.99995, \\ gap &= x^T p + v^T s + t_1^T z_1 + t_2^T z_2, \\ \gamma &= \sigma \frac{gap}{2(N + TP)}, \quad \sigma = \frac{1}{2\sqrt{N + TP}}. \end{aligned} \tag{4.32}$$

Algorithm 4 – Primal-Dual Interior-Point Algorithm to solve the radiotherapy problem (4.18)

Data: Variables: $x^0, \xi^0, v^0, {z_1}^0, {z_2}^0, {t_1}^0, {t_2}^0, {w_1}^0, {w_2}^0$
Parameters: $A, \tilde{b}, U, \tau, \sigma, tol$
begin
  while $CP > tol$ && $k < maxit$ **do**
    Compute $\mu(\xi)$ and $s(\xi)$;
    Compute ∇f and H;
    Compute gap and γ;
    Compute the residues:;
    $r_1 = A^T q + p + y - \nabla f(x)$;
    $r_2 = w_1 + w_2 - q$;
    $r_3 = -y - s$;
    $r_4 = t_1 - w_1$;
    $r_5 = -t_2 - w_2$;
    $r_6 = U - x - v$;
    $r_7 = b_1 + z_1 - \xi$;
    $r_8 = b_2 - z_2 - \xi$;
    $r_9 = \xi - Ax$;
    $r_{10} = \gamma e - XPe$;
    $r_{11} = \gamma e - VSe$;
    $r_{12} = \gamma e - T_1 Z_1 e$;
    $r_{13} = \gamma e - T_2 Z_2 e$;
    Compute the directions: dx in (4.26) and the other from system (4.24);
    Determinate α_p and α_d using (4.28) and (4.29), respectively;
    Update the variables;
    $x^{k+1} = x^k + \alpha_p dx$;
    $\xi^{k+1} = \xi^k + \alpha_p d\xi$;
    $v^{k+1} = v^k + \alpha_p dv$;
    $z_1^{k+1} = z_1^k + \alpha_p dz_1$;
    $z_2^{k+1} = z_2^k + \alpha_p dz_2$;
    $p^{k+1} = p^k + \alpha_d dp$;
    $s^{k+1} = s^k + \alpha_d ds$;
    $t_1^{k+1} = t_1^k + \alpha_d dt_1$;
    $t_2^{k+1} = t_2^k + \alpha_d dt_2$;
    $y^{k+1} = y^k + \alpha_d dy$;
    $q^{k+1} = q^k + \alpha_d dq$;
    $w_1^{k+1} = w_1^k + \alpha_d dw_1$;
    $w_2^{k+1} = w_2^k + \alpha_d dw_2$;
    Compute CP from (4.31)
  end
end

5 Computational Results

Algorithm 4 is implemented in MATLAB Version 7.10.0.499 ($R2010a$) on a computer with characteristics: Windows 10 Pro Education, Intel(R) Core(TM) i7-4790 CPU, 3.60GHz and 16GB of RAM. In this chapter we present the results obtained using the set of instances *Head-and-Neck* (6.5GiB), from the TROTS dataset[1], that consists of 15 patients with cancer in the head-and-neck region. More details and explanation about the contents and how to access to these data are in [6].

5.1 Matrix Properties

The set of instances *Head-and-Neck* (6.5GiB), from the TROTS dataset [6], provides the numerical data parameters as the quantity of beamlets used on each patient; pencil-beam dose matrices for each clinical structures where the number of rows indicates the number of voxels for each structure; dose values and a sufficient value indicating that the dose value not need to become lower than the given sufficient value (we will use these values for limiting the tumor dose values). All these parameters and data are summarized in the following tables.

Table 1 presents the quantity of voxels (*columns* TP) used to represent each structure and the sparsity of them (*columns* SP) being the percentage of elements non-zero. For our experiments we took the matrix of the tumor (T), A_T, and just 3 critical organs near to the tumor region: Spinal Cord (C_1), A_{C_1}; Brain-stem (C_2), A_{C_2}, and Larynx (C_3), A_{C_3}.

It can be seen that A_{C_2} is much more sparse than the other structures, where it is null for the Patient 5. A_T and A_{C_3} are a little dense, except for the Patient 5 which are considered full matrices. A_{C_3} is slightly less sparse than A_T, except for the Patient 7.

Table 2 describe the properties of the general matrix constituted by the tumor and critical organs matrices, that is, $A = (A_T \;\; A_{C_1} \;\; A_{C_2} \;\; A_{C_3})^T$.The *column* 2 provides the total of beamlets used N, representing the total number of variables (the number of columns). *Column* 3 has the total number of voxels TP, and *column* 4, the sparsity of the resulting matrix A. We using the *rank* function of MATLAB to compute the rank of A and highlight that A has full rank, for all patients.

The expected values of doses (in Gray - Gy) for the structures considered are the same for all the patients [7], and are presented in Table 3. For the critical organs, these

[1] Link access: https://www.erasmusmc.nl/en/cancer-institute/research/projects/radiotherapy-trots. (Last access date: 16-08-2020)

Table 1 – Properties of the pencil-beam matrix of each structure: quantity of voxels TP and sparsity SP. A_T matrix of the Tumor, A_{C_1} matrix of the Spinal cord, A_{C_2} matrix of the Brain-stem and A_{C_3} matrix of the Larynx.

Patient	A_T		A_{C_1}		A_{C_2}		A_{C_3}	
	TP	$SP(\%)$	TP	$SP(\%)$	TP	$SP(\%)$	TP	$SP(\%)$
1	5 096	17.98	3 529	13.30	3 757	4.92	5 263	19.72
2	5 167	26.87	3 181	18.85	3 126	1.80	5 075	30.58
3	5 176	24.93	3 065	15.87	3 397	1.19	5 127	29.07
4	5 120	17.63	3 775	12.94	3 137	3.14	5 339	19.73
5	5 208	70.59	2 068	18.84	2 101	0.00	3 615	76.33
6	5 154	25.98	2 891	17.24	3 046	2.99	5 011	30.16
7	5 110	17.30	3 603	12.19	3 500	2.20	5 134	16.85
8	5 110	17.28	3 429	12.02	4 133	6.41	5 142	18.66
9	5 126	19.97	3 906	12.82	3 920	1.20	5 131	21.43
10	5 127	17.15	3 406	12.42	3 674	6.03	5 105	18.44
11	5 128	18.98	3 660	11.70	3 567	1.29	5 159	20.63
12	5 162	22.75	3 290	14.94	3 853	1.70	5 247	25.18
13	5 107	18.36	3 279	11.70	3 669	4.92	5 101	20.17
14	5 108	20.05	3 993	12.78	3 309	5.56	3 970	22.19
15	5 113	21.17	3 213	14.68	3 580	1.17	5 238	24.06

doses are the maximum allowed value.

However, for our problem, we suggest to use *triangular-fuzzy-numbers*. Thus, we be inclined to think that the tumor cells can receive a little more radiation and need a minimal radiation to be damaged, according to the sufficient value from the dataset [7], we take the fuzzy target as $\tilde{b}_T = (0.5d_T \;\; d_T \;\; 1.1d_T)$. In other words, a dose $b_T = 48.3$ kills the tumor cells and the values $b_{T1} = 24.15$ is sufficient lower bound and $b_{T2} = 53.13$ are reasonable upper bound. These values are reasonable limits when taking account the sufficient dose value and increasing the upper bounds can make an increasing in the dose values for the surrounding structures.

Unlike the tumor, critical organs can not receive more radiation than their limits d_{C_i}, $i = 1, 2, 3$; and the lower the radiation the lower damage for these organs. Thus, after having tested with different percentages of d_{C_i}, the fuzzy target for each $i = 1, 2, 3$ is $\tilde{b}_{C_i} = (0 \;\; 0.3d_{C_i} \;\; d_{C_i})$, where $30\%d_{C_i}$ allows us to have a good performance of the algorithm. The values of this analyze are summarized in Table 4.

5.2 Computing the Initial Point

The main goal of the radiotherapy problem is to eliminate the tumor cells since without the requirement of a minimal dose to kill tumor cells, the optimal delivered radiation for the another cells will be zero. For this reason, the choice of the initial delivered

Table 2 – Properties of the resulting matrix $A = (A_T \;\; A_{C_1} \;\; A_{C_2} \;\; A_{C_3})^T$: dimension $TP \times N$ and sparsity.

Patient	Total of Voxels(TP)	Total of Beamlets(N)	Total sparsity(%)
1	17 645	9 977	14.78
2	16 549	6 331	21.73
3	16 765	6 939	19.73
4	17 371	10 125	14.64
5	12 992	1 815	52.53
6	16 102	6 690	21.36
7	17 347	9 646	13.06
8	17 814	10 299	14.14
9	18 083	8 563	14.77
10	17 312	10 579	14.24
11	17 514	9 184	14.34
12	17 552	7 233	17.39
13	17 156	9 784	14.75
14	16 380	8 940	15.87
15	17 144	8 431	16.66

Table 3 – Dose values for the Tumor (T), Spinal cord (C_1), Brain-stem (C_2) and Larynx (C_3).

Structure	Dose d (Gy)
T	48.30
C_1	38.00
C_2	38.00
C_3	48.30

Table 4 – Triangular-fuzzy-number dose $\tilde{b} = (b_1 \;\; b \;\; b_2)$ for each structure: Tumor (T), Spinal cord (C_1), Brain-stem (C_2) and Larynx (C_3).

Structure	b_1 (Gy)	b (Gy)	b_2 (Gy)
T	24.15	48.30	53.13
C_1	0.00	11.40	38.00
C_2	0.00	11.40	38.00
C_3	0.00	14.49	48.30

radiation, x^0, will be based on the desired dose of the tumor.

Only pencil-beams that hit at least one tumor voxel is considered, that is, the matrix to the tumor. Then the algorithm starts the process of minimizing the delivered radiation to the other organs and tissues considering that the tumor cells have a minimal dose that must be satisfied.

The data set [7] also suggests use an **InitialiseMatrix**, matrix required for initialising the problem, in this case is tumor matrix A_T; an **InitialiseReferenceDose**, the reference dose d, which is the right-hand side for the least-square $A_T x = d$; and furthermore a **InitialiseRegularisatioMatrix**, is a smoothing matrix B [6] (symmetric and positive definite) that helps to regulate the least-square problem (the use of this matrix is justifying in [5]). Therefore, we compute x^0 by solving the following least-square problem:

$$\begin{pmatrix} A_T \\ B \end{pmatrix} x^0 = \begin{pmatrix} d \\ 0 \end{pmatrix}. \tag{5.1}$$

We used the *lsqr*-function from MATLAB and the solutions obtained have relative residual ≈ 0.01.

The radiation delivered inside each structure with x^0 is presented in the Table 5. As it can be seen, the tumor receives high radiation dose and, at the same time, the Larynx C_3 is receiving almost the limiting value causing potential damage to this structure. Although the other structures, C_1 and C_2, are not receiving high dose, these values can be improved by Algorithm 4.

5.3 Numerical experiments

Table 6 presents the results obtained using the parameters 4.32. Furthermore, analysing the radiation flow from the solution provides by [7], we establishing an upper bound for the variable x as $U = 1000$. All the instances stopped after the maximum number of iteration was attained without achieving the desirable precision, namely $CP < tol$. In the objective-function values (4.2), in column FO, the minimum value was for Patient 7 due to the percentage of sparsity and, as A is a full matrix for Patient 5 the objective function has a high value.

It was difficult to obtain a large decrease in the amount of radiation in this region, since both of the tumor and larynx matrices (A_T and A_{C_3}) have similar percentage of sparsity. This can be seen by looking at the large the surprise value (4.1), see Table 7. For this reason, the objective function values are high. However, this does not mean that the solution obtained, x_{sol}, is not good since, despite not receiving the "ideal value" b^{C_3} for the fuzzy number, the radiation deposited inside the larynx region was, in the best case, for Patient 12, of 0.333%, and for the worst case, Patient 5, of 0.598%, of the allowed maximum value d_{C_3}. That is, instead of the variable ξ taking on values around of 14.49 (and thus $s(\xi) \approx 0$), ξ is, on average, approximately 25. This dose will not significantly harm the larynx, according to the radiation oncologist [6], Table 3.

The value of the delivered radiation of each beamlet, x_{sol}, is such that the average of radiation dose received by the spinal cord and the brain-stem, Ax_{sol}, is better

Table 5 – Radiation delivered inside each structure with the initial solution x^0

Patient	S	Ax_0 (Gy)
1	T	47.1905
	C1	37.3800
	C2	6.8536
	C3	47.1290
2	T	47.2073
	C1	22.8167
	C2	0.5606
	C3	47.2424
3	T	47.0911
	C1	20.5297
	C2	0.2476
	C3	47.0412
4	T	47.1437
	C1	23.8770
	C2	4.3842
	C3	46.6533
5	T	47.1443
	C1	7.5957
	C2	0.0000
	C3	47.1289
6	T	47.1053
	C1	22.9271
	C2	1.3798
	C3	47.0482
7	T	47.1439
	C1	19.8742
	C2	1.9662
	C3	39.4022
8	T	47.1451
	C1	22.2247
	C2	11.1697
	C3	43.0122

Patient	S	Ax_{sol} (Gy)
9	T	47.0945
	C1	18.8092
	C2	0.3779
	C3	43.1636
10	T	47.1416
	C1	23.5529
	C2	10.0669
	C3	46.6737
11	T	47.1418
	C1	19.7682
	C2	0.5949
	C3	43.4228
12	T	47.1442
	C1	20.1333
	C2	0.4682
	C3	42.9688
13	T	47.1918
	C1	20.8242
	C2	7.1040
	C3	44.2318
14	T	47.2034
	C1	20.6710
	C2	7.0418
	C3	45.1990
15	T	47.1382
	C1	24.6830
	C2	0.4257
	C3	47.1687

than ξ (see columns Ax_{sol} in Table 8 and ξ in Table 7), where for most of them close to zero. On these structures the difference $Ax_{solution} - \xi$ increases the value of the residue r_9 and then according to (4.30b), also increases the value of *gap*.

It is not possible to make direct comparisons between the solution obtained by our algorithm, x_{sol}, and the optimal solution provided in [6], x^* (obtained by the Erasmus-iCycle solver developed in [8]), because they consider all the critical organs and healthy tissue in the objective function and it is optimized taking into account different levels of priority to attack each region. However, we are able to compare the radiation delivered, by x_{sol} and x^*, inside each structure computing the radiation dose received, Ax_{sol} and Ax^*. In the Table 8 we present the average of these values for each region

Table 6 – Results for all patients considering the tumor, spinal cord, brain-stem and larynx structures, with FO value of the objective function (4.2) and the *time* is in seconds.

Patient	FO	gap	$time(s)$
1	774.115	64.87	196750.5
2	828.610	1439.80	141215.7
3	810.823	1172.70	167628.8
4	628.447	123.57	197510.1
5	3931.616	4929.10	48124.2
6	858.412	1514.70	152091.8
7	76.881	667.62	152252.9
8	287.457	623.23	190976.3
9	281.043	1268.80	164535.6
10	621.642	279.46	199699.8
11	193.020	979.80	170051.7
12	81.178	827.04	153251.8
13	231.003	446.75	186811.9
14	260.587	539.78	172100.6
15	825.580	648.42	167732.3

(Tumor T, Spinal cord C_1, Brain-stem C_2 and Larynx C_3) and patient.

The radiation dose received by the tumor with the solution x_{sol}, in the worst case, represents 69.58% of the expected dose and it is 39.16% higher than the desired minimal dose. For all patients, the radiation delivered by x_{sol} is lower than x^*. However, it is still higher than the sufficient value given by the oncologist in order to shrink the tumor [7]. Moreover, the radiation delivered by x_{sol} inside the Spinal-cord is smaller than 29% of the maximum tolerance dose; for the Brain-stem, it is smaller than 18% of the tolerance; and, since the Larynx-matrix is less sparse than the other structures, the radiation is lower than 59.77% of the tolerance. Thus, on the critical structures, unlike x^*, our solution x_{sol} delivers a little amount of radiation for most of the patients, helping to maintain theses structures healthy.

Table 7 – Average for each structure (Tumor T, Spinal cord C_1, Brain-stem C_2 and Larynx C_3) of the radiation ξ received, membership function μ (4.10) and surprise function $\mathcal{S}(\xi)$ (4.1).

Patient	S	ξ	μ	$\mathcal{S}(\xi)$	Patient	S	ξ	μ	$\mathcal{S}(\xi)$
1	T	47.9725	0.9854	0.0003	9	T	48.0595	0.9901	0.0001
	C1	12.6161	0.9727	0.0026		C1	11.7636	0.9810	0.0009
	C2	11.7039	0.9886	0.0012		C2	11.4018	0.9999	0.0000
	C3	23.3933	0.7368	0.1379		C3	18.3844	0.8849	0.0503
2	T	48.0508	0.9877	0.0002	10	T	47.9379	0.9846	0.0004
	C1	12.5338	0.9575	0.0049		C1	13.1878	0.9329	0.0135
	C2	11.4036	0.9999	0.0000		C2	11.9342	0.9800	0.0027
	C3	24.0058	0.7187	0.1556		C3	21.9280	0.7801	0.1061
3	T	48.0879	0.9905	0.0001	11	T	48.0842	0.9912	0.0001
	C1	11.9898	0.9779	0.0014		C1	11.7014	0.9865	0.0005
	C2	11.4016	0.9999	0.0000		C2	11.4013	0.9999	0.0000
	C3	23.9330	0.7280	0.1530		C3	17.4034	0.9139	0.0333
4	T	47.9584	0.9842	0.004	12	T	48.0395	0.9893	0.0002
	C1	12.7778	0.9483	0.0059		C1	11.4802	0.9885	0.0003
	C2	11.4490	0.9982	0.0000		C2	11.4016	0.9999	0.0000
	C3	22.1117	0.7747	0.1090		C3	16.0917	0.9527	0.0115
5	T	45.9931	0.9042	0.1423	13	T	47.9883	0.9868	0.0002
	C1	11.7435	0.9871	0.0011		C1	12.2714	0.9673	0.0028
	C2	11.4004	1.0000	0.0000		C2	11.5688	0.9937	0.0004
	C3	28.5401	0.5845	0.8325		C3	18.1496	0.8919	0.0389
6	T	48.0548	0.9881	0.0002	14	T	47.9710	0.9861	0.0002
	C1	12.6037	0.9548	0.0056		C1	12.4373	0.9611	0.0040
	C2	11.4111	0.9996	0.0000		C2	11.5685	0.9937	0.0004
	C3	24.2069	0.7127	0.1633		C3	19.2540	0.8592	0.0557
7	T	47.9169	0.9842	0.0003	15	T	48.0470	0.9887	0.0002
	C1	12.0071	0.9773	0.0014		C1	11.9510	0.9794	0.0009
	C2	11.4039	0.9999	0.0000		C2	11.4028	0.9999	0.0000
	C3	16.3250	0.9458	0.0098		C3	23.9703	0.7197	0.1529
8	T	47.4940	0.9667	0.0024					
	C1	13.5550	0.9192	0.0199					
	C2	12.1525	0.9718	0.0062					
	C3	18.8855	0.8701	0.0311					

Finally, in order to compare radiation dose received in the structures by the obtained plan, we compute the percentage of absorbed dose for each structure volume and show graphically in Histograms Dose - Volume[2]. Ideally, the entire tumor region must be receive 100% of the prescribe dose (Table 3) and the other structure 0% of radiation, but it is impossible. For the developed algorithm we have the next representation of absorbed dose.

[2] A dose-volume histogram (DVH) summarizes the simulated radiation distribution within a volume of interest of a patient which would result from a proposed radiation treatment plan [12]

Table 8 – Comparing of the radiation delivered inside each structure with our solution x_{sol} and the solution x^* provided in [7] (obtained by the Erasmus-iCycle solver [8])

Patient	S	Ax_{sol} (Gy)	Ax^* (Gy)
1	T	41.0216	47.4951
	C1	10.4115	34.9725
	C2	3.5683	4.7520
	C3	23.5570	47.1370
2	T	40.5938	47.5038
	C1	9.6054	17.0301
	C2	0.7316	0.4131
	C3	24.1365	47.6418
3	T	39.7145	47.4923
	C1	8.6460	13.5169
	C2	0.2965	0.2001
	C3	24.0148	47.4134
4	T	41.6663	47.4868
	C1	10.8820	18.2829
	C2	3.1957	2.6366
	C3	22.3461	41.6452
5	T	33.6077	47.4181
	C1	4.9491	8.2195
	C2	0.0000	0.0000
	C3	28.8702	47.9382
6	T	40.7933	47.5048
	C1	10.5641	17.6712
	C2	1.1941	0.9897
	C3	24.4033	47.9296
7	T	42.3733	47.5196
	C1	9.5406	15.1842
	C2	2.2035	1.6192
	C3	16.4151	33.3597
8	T	42.2304	47.5468
	C1	10.8487	14.3067
	C2	6.5405	9.0553
	C3	19.2812	34.5001
9	T	40.6694	47.5346
	C1	8.7181	13.0790
	C2	0.6271	0.2205
	C3	18.3803	14.5001
10	T	41.9435	47.5415
	C1	10.7642	15.8300
	C2	5.1489	6.7971
	C3	22.1383	43.0029
11	T	41.3433	47.5292
	C1	9.0392	14.0455
	C2	0.5194	0.4483
	C3	17.4055	34.5001
12	T	41.5919	47.4920
	C1	8.6512	14.0062
	C2	0.4630	0.2833
	C3	16.0901	34.5001
13	T	42.2915	47.5369
	C1	8.6080	13.6462
	C2	4.0750	6.5379
	C3	18.2525	34.5000
14	T	42.2790	47.5085
	C1	8.6806	14.3514
	C2	4.2887	6.3614
	C3	19.3913	34.5000
15	T	39.9717	47.5815
	C1	9.9673	17.0237
	C2	0.5934	0.2875
	C3	24.0366	47.8879

Patient 1 Figure 6, 75% of the tumor volume receives 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 20% of the Spinal cord C_1 volume receives between $39\% - 52\%$ of maximal prescribe dose; less than 25% of the Brain-stem C_2 volume receives between $13\% - 39\%$ of maximal prescribe dose; and less than 10% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

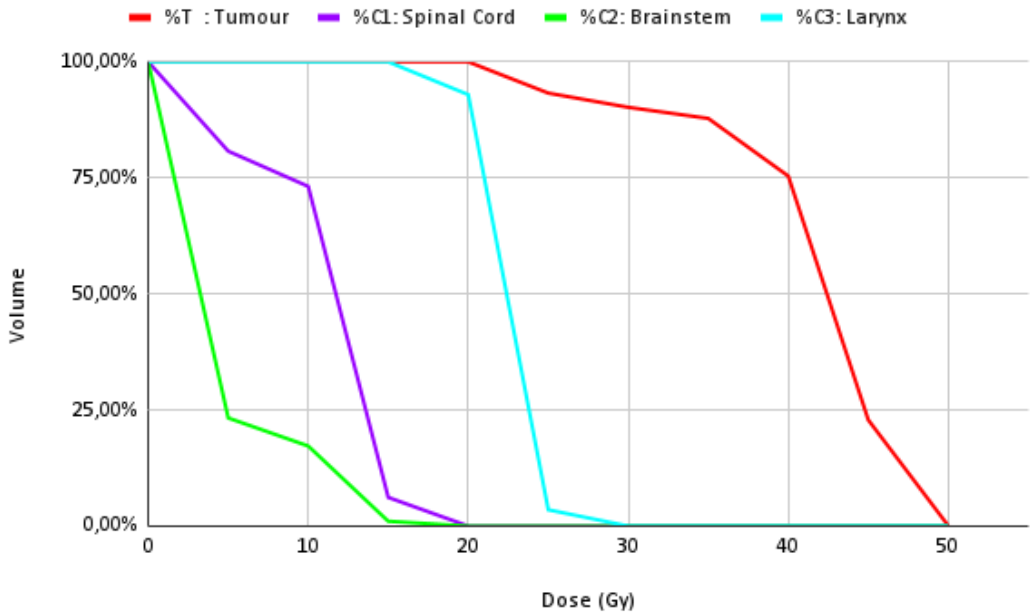


Figure 6 – Patient 1

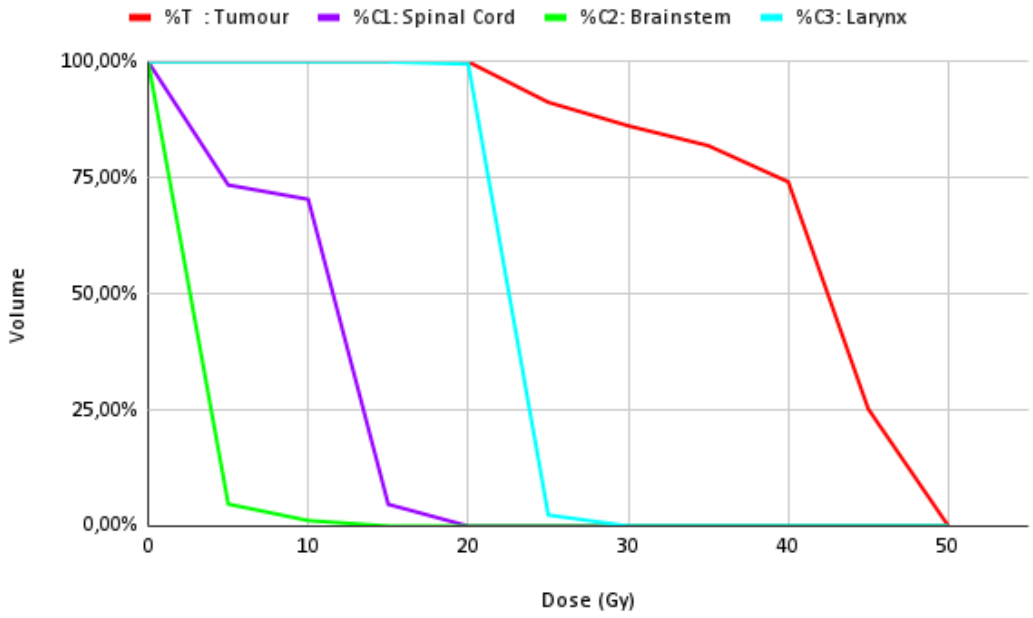


Figure 7 – Patient 2

Patient 2 Figure 7, 75% of the tumor volume receives 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 20% of the Spinal cord C_1 volume receives between $39\% - 52\%$ of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives between $13\% - 39\%$ of maximal prescribe dose; and less than 10% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

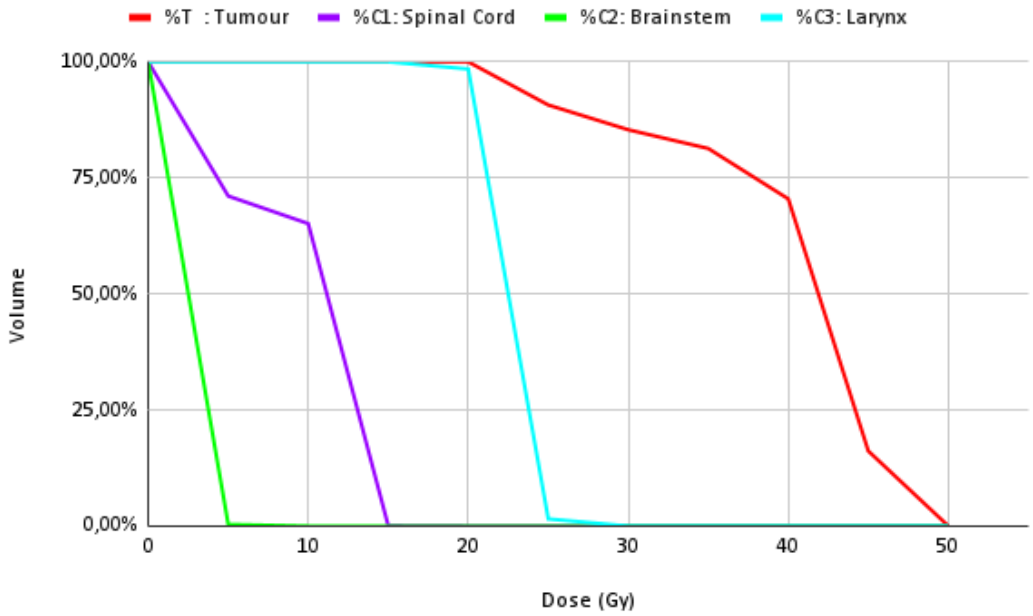


Figure 8 – Patient 3

Patient 3 Figure 8, approximately 75% of the tumor volume receives 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 50% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 25% of the Brain-stem C_2 volume receives less than 13% of maximal prescribe dose; and less than 25% of the Larynx C_3 volume receives less than 51% of maximal prescribe dose.

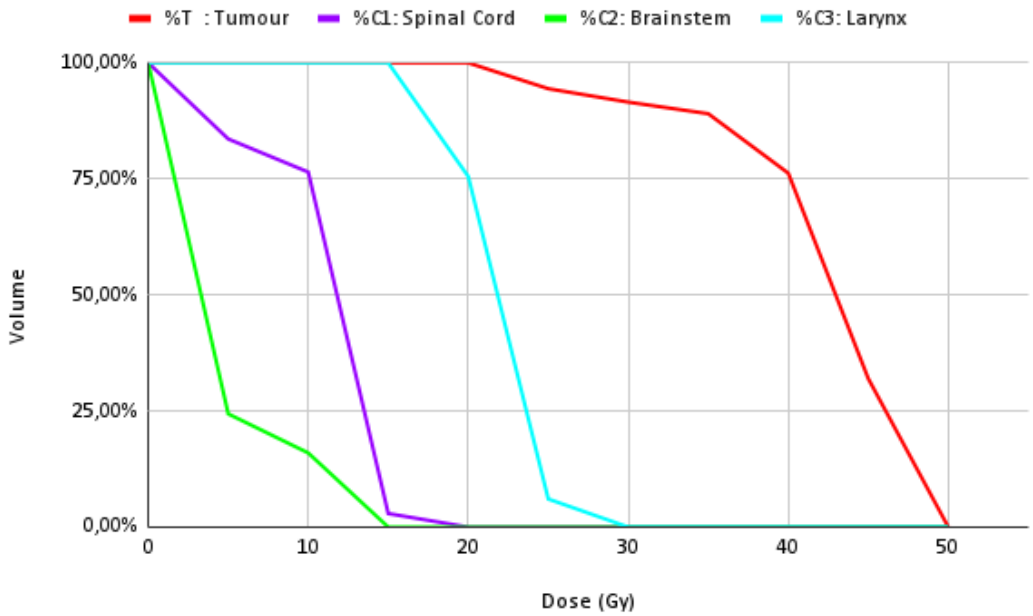


Figure 9 – Patient 4

Patient 4 Figure 9, 75% of the tumor volume receives 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 10% of the Spinal cord C_1 volume receives less than 52% of maximal prescribe dose; less than 20% of the Brain-stem C_2 volume receives less than 39% of maximal prescribe dose; and less than 20% of the Larynx C_3 volume receives between 62% of maximal prescribe dose.

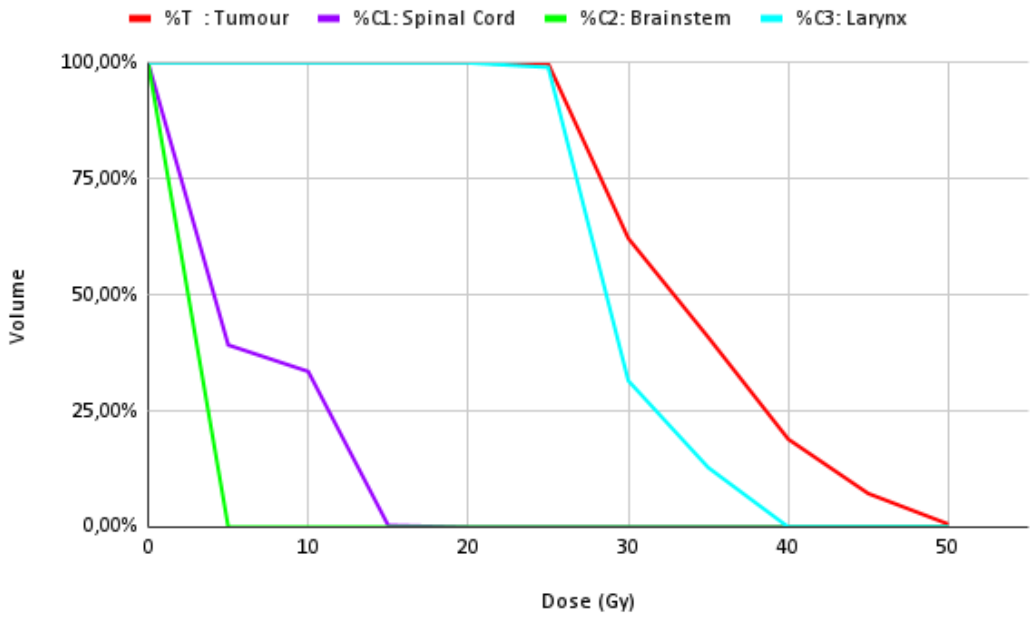


Figure 10 – Patient 5

Patient 5 Figure 10, 75% of the tumor volume receives 30 Gy equivalent to almost 62% of the ideal value for the oncologist and 25% between $82\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 25% of the Brain-stem C_2 volume receives between 13% of maximal prescribe dose; and almost 100% of the Larynx C_3 volume receives 51% of maximal prescribe dose.

Patient 6 Figure 11, 75% of the tumor volume receives 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 20% of the Spinal cord C_1 volume receives less than 52% of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives between $13\% - 39\%$ of maximal prescribe dose; and less than 10% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

Patient 7 Figure 12, more than 75% of the tumor volume receives 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives less than 39%

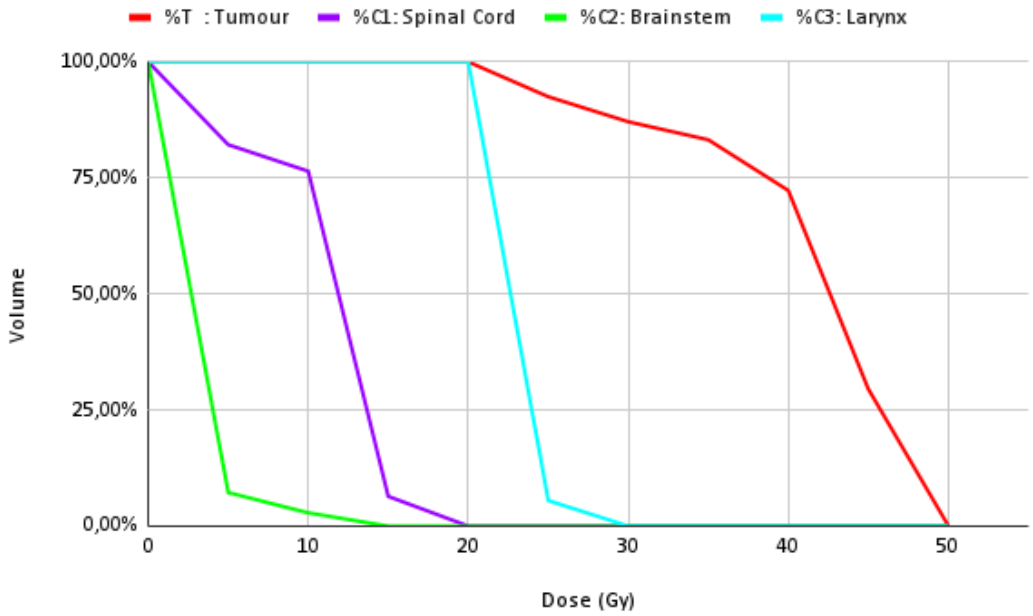


Figure 11 – Patient 6

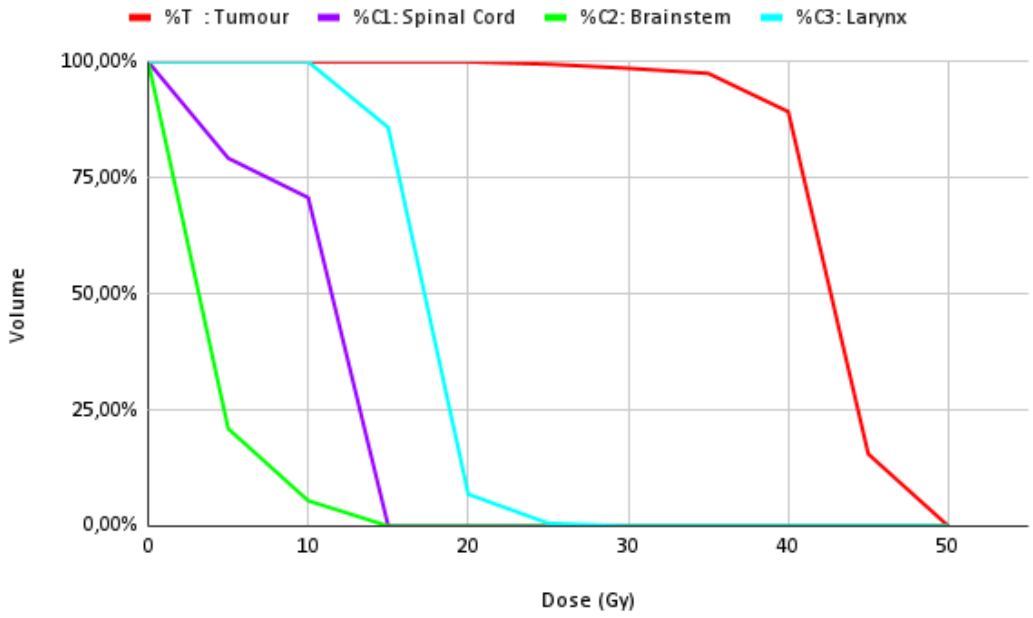


Figure 12 – Patient 7

of maximal prescribe dose; and less than 15% of the Larynx C_3 volume receives less than 51% of maximal prescribe dose.

Patient 8 Figure 13, more than 75% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives

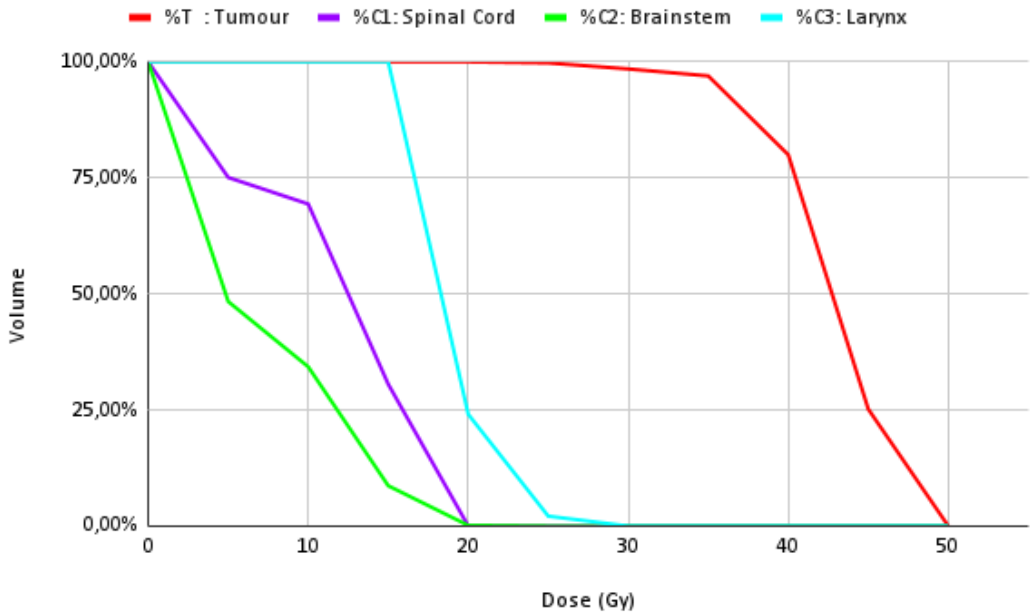


Figure 13 – Patient 8

between $39\% - 52\%$ of maximal prescribe dose; less than 20% of the Brain-stem C_2 volume receives between $39\% - 52\%$ of maximal prescribe dose; and less than 15% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

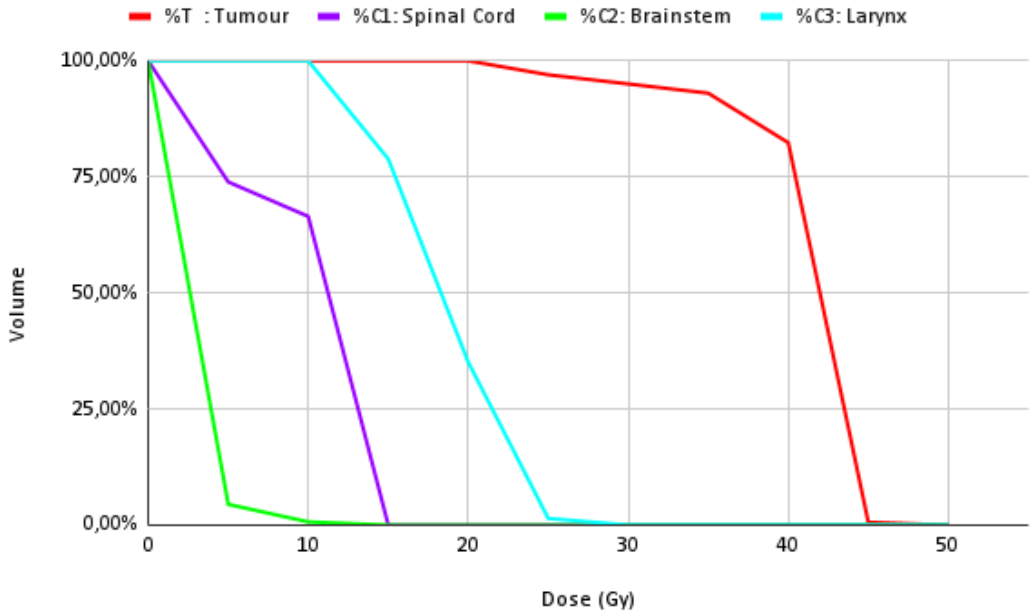


Figure 14 – Patient 9

Patient 9 Figure 14, 80% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% approximately 93%

of the ideal value. Less than 25% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives less than 26% of maximal prescribe dose; and less than 10% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

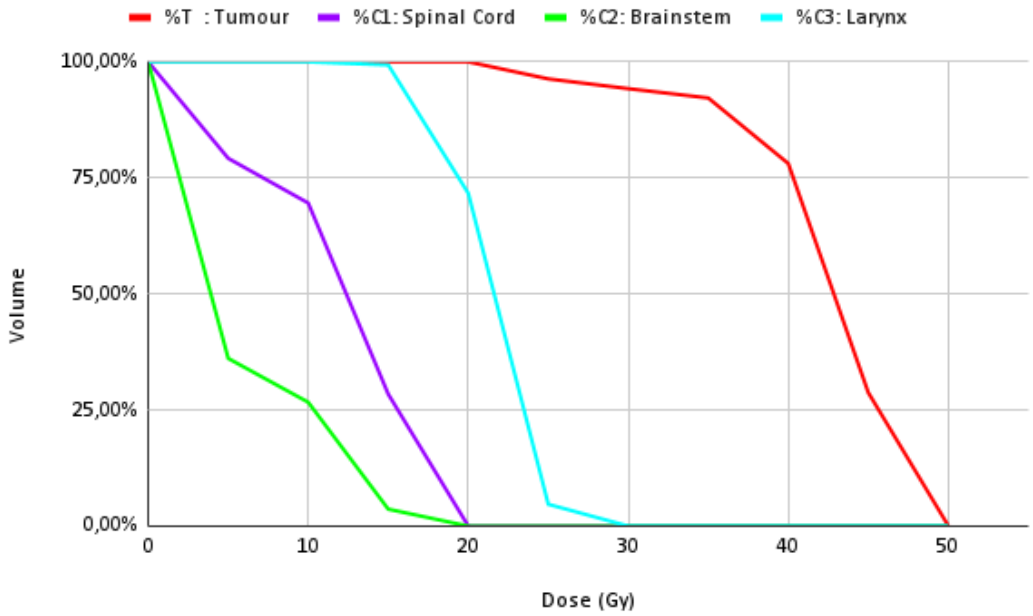


Figure 15 – Patient 10

Patient 10 Figure 15, more than 80% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives between $39\% - 52\%$ of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives between $39\% - 52\%$ of maximal prescribe dose; and less than 10% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

Patient 11 Figure 16, more than 80% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives between $13\% - 26\%$ of maximal prescribe dose; and less than 25% of the Larynx C_3 volume receives between $40\% - 51\%$ of maximal prescribe dose.

Patient 12 Figure 17, more than 80% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 10% of the Brain-stem C_2 volume receives between $13\% - 26\%$ of maximal prescribe dose; and less than 20% of the Larynx C_3 volume receives between $40\% - 51\%$ of maximal prescribe dose.

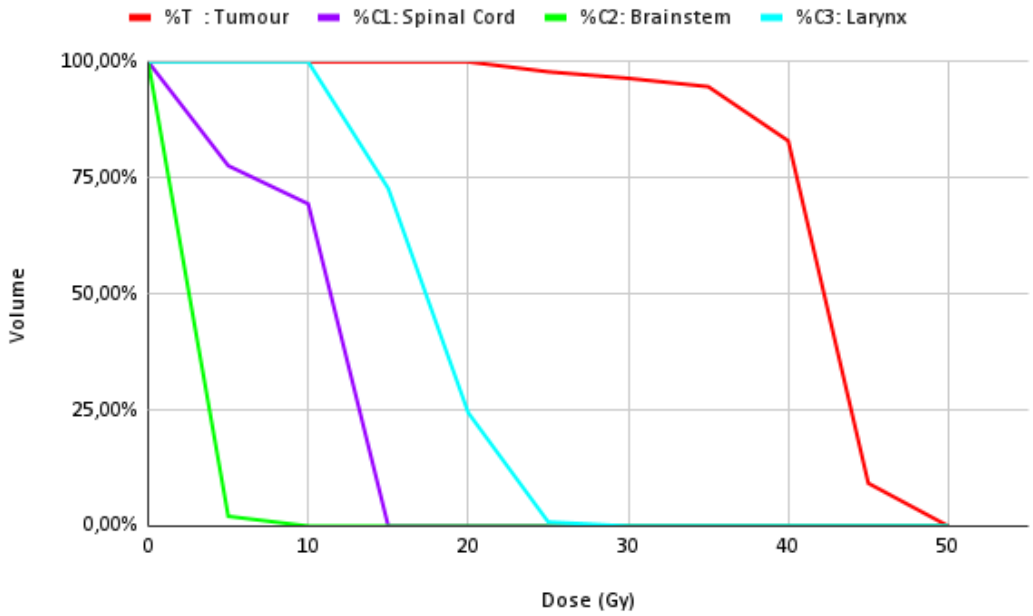


Figure 16 – Patient 11

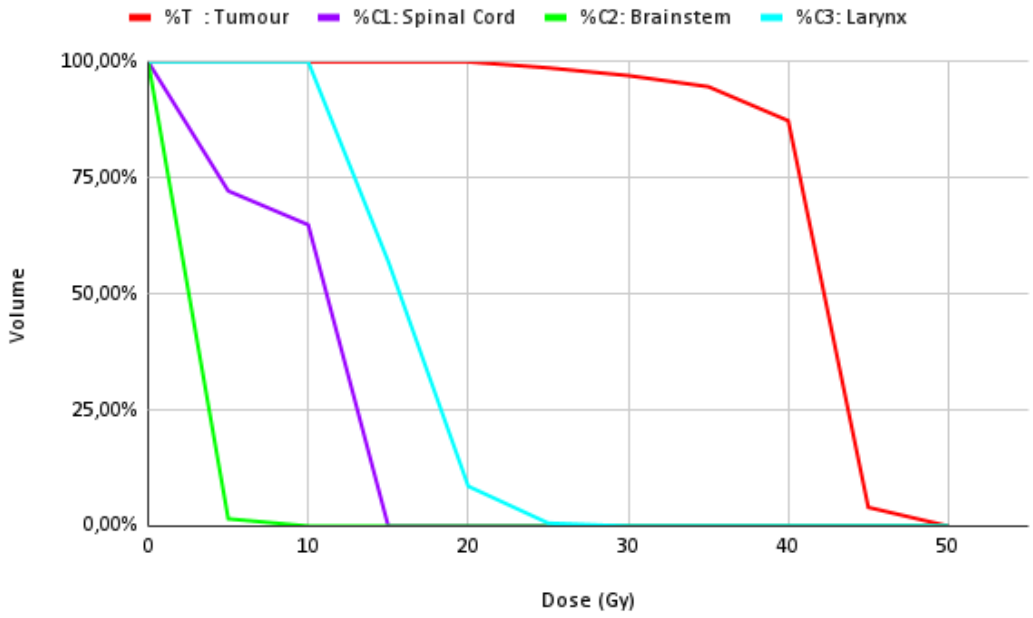


Figure 17 – Patient 12

Patient 13 Figure 18, 80% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 and the Brain-stem C_2 volume receives between $26\% - 39\%$ of maximal prescribe dose; and less than 25% of the Larynx C_3 volume receives between $40\% - 51\%$ of maximal prescribe dose.

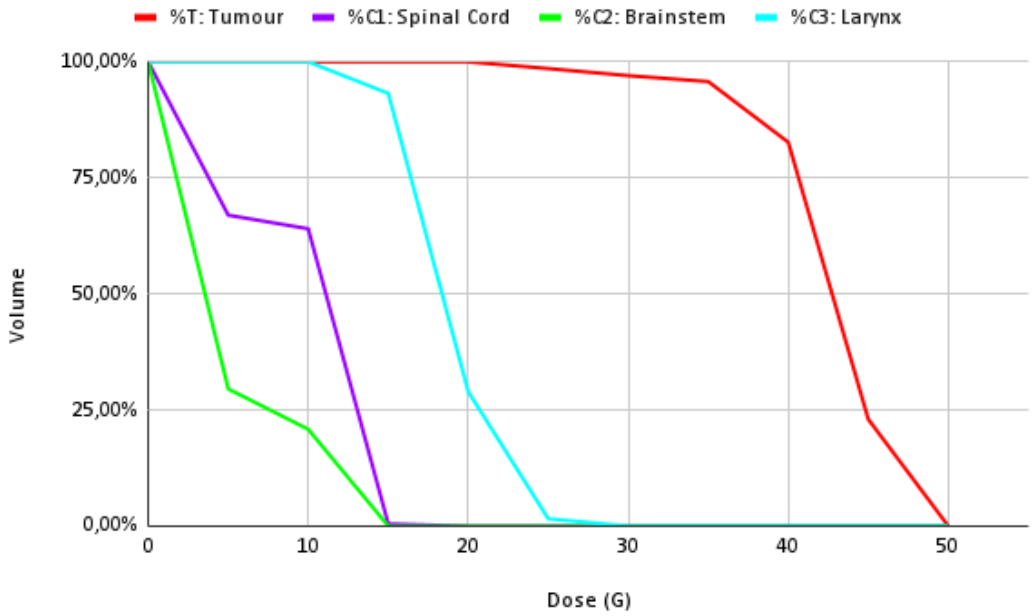


Figure 18 – Patient 13

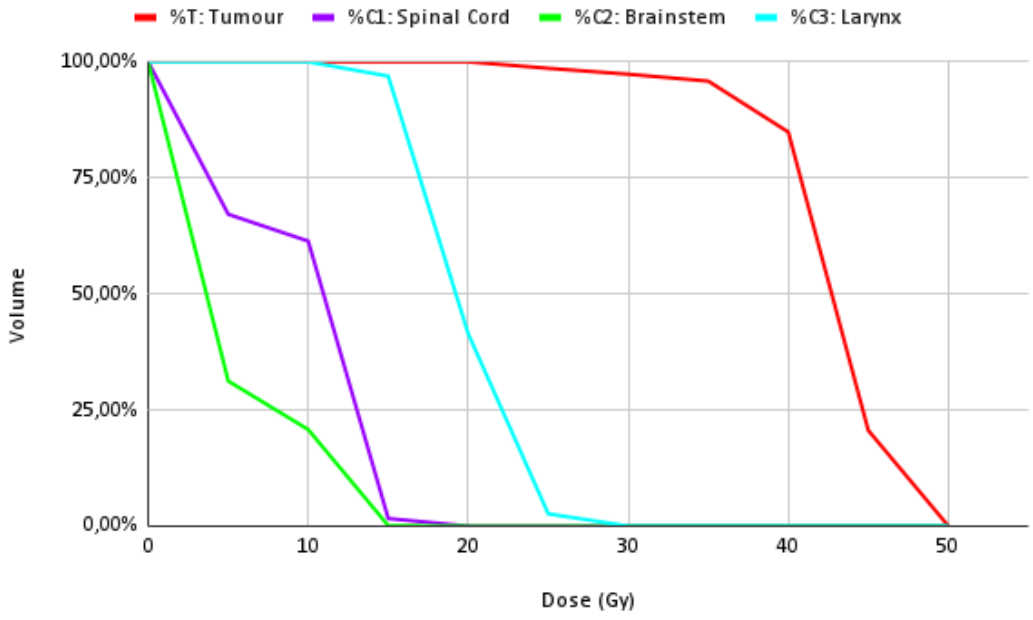


Figure 19 – Patient 14

Patient 14 Figure 19, 80% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and 25% between $93\% - 103\%$ of the ideal value. Less than 25% of the Spinal cord C_1 volume receives less than 39% of maximal prescribe dose; less than 25% of the Brain-stem C_2 volume receives between $26\% - 39\%$ of maximal prescribe dose; and less than 25% of the Larynx C_3 volume receives between $40\% - 51\%$ of maximal prescribe dose.

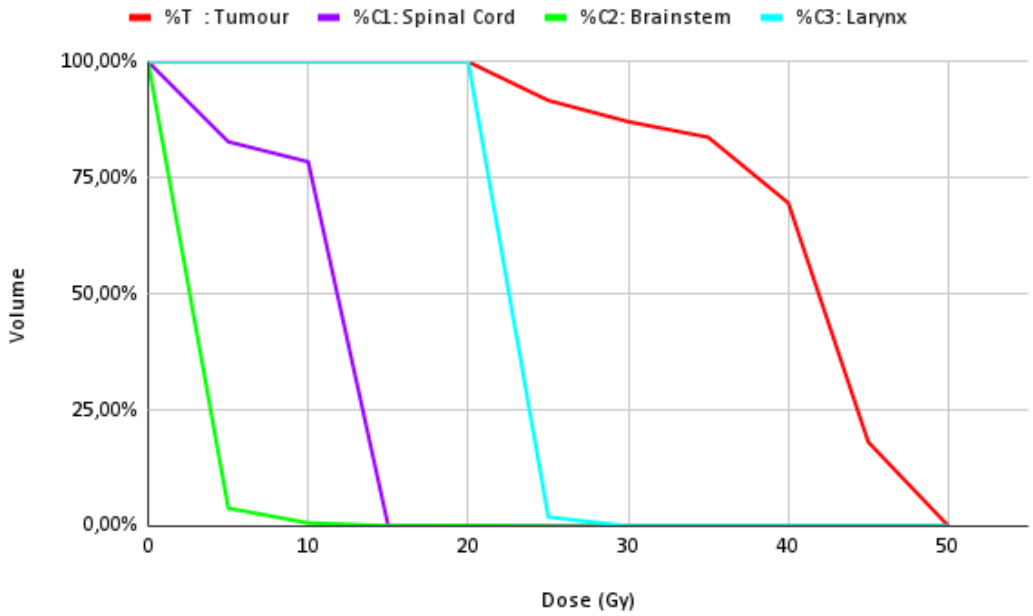


Figure 20 – Patient 15

Patient 15 Figure 20, more than 70% of the tumor volume receives more than 40 Gy equivalent to almost 83% of the ideal value for the oncologist and approximately 25% between $93\% - 103\%$ of the ideal value. Less than 10% of the Spinal cord C_1 volume receives between $13\% - 26\%$ of maximal prescribe dose; less than 25% of the Brain-stem C_2 volume receives less than 39% of maximal prescribe dose; and less than 10% of the Larynx C_3 volume receives between $51\% - 62\%$ of maximal prescribe dose.

Also, less than 25% critical structures volume receive less than 40% of allowed tolerance dose, and in the worst case, approximately 10% of Larynx volume receive between $51\% - 62\%$ of allowed tolerance dose, thus these solutions meet the expectations of maintaining their functionality. In most of patients approximately 100% of the tumor volume receives the sufficient value [6], and more than 80% of the tumor volume receives more than 80% of the ideal value for the oncologist [6].

6 Conclusions and Future Proposals

This research developed a primal-dual interior-point algorithm to solve the dose distribution problem of the radiation therapy treatment planning using the surprise function approach as a measure of violation of the linear constraints, whose dose bounds were triangular fuzzy numbers, leading to a non-linear objective function. Then we tested the algorithm on the Head-and-Neck dataset (provided by [6]). The analyse of the numerical experiments allows us to conclude that the proposed algorithm is able to solve large-scale problems finding satisfactory solutions. These solutions allow the tumor to receive a radiation higher than the value given by the oncologist to shrink it. The proposed treatment plans are good attacking the tumor, since they deliver radiation very close to the oncologist prescribed doses. Furthermore, by delivering doses within the acceptable limits they preserve the healthy tissues and organs at risk, avoiding unacceptable damage to such tissues. Furthermore, when we compare with the solution obtained by iCycle [6], our solution in addition to attacking the tumor and shrinking it, delivers less amount of radiation to the critical structures, guaranteeing a good treatment plan.

Future proposal is to enhance the algorithm and its implementation in order to obtain solution in less computational time, and thus test it with other large datasets. With respect to the implementation it will be useful to do it in languages as C or FORTRAN. Furthermore, it will be interesting to study and compare situations considering another types of membership functions and fuzzy numbers.

Bibliography

[1] Bahr, G. K., Kereiakes, J. G., Horwitz, H., Finney, R., Galvin, J., and Goode, K. The method of linear programming applied to radiation treatment planning. *Radiology 91*, 4 (1999), 686 – 693. pages 17

[2] Bazaraa, M. S.; Jarvis, J. J., and Sherali, H. D. *Linear Programming and Network Flows*, 4 ed. John Wiley Sons, 2010. pages 21, 22

[3] Bertsekas, D. P. *Nonlinear Programming*, 2 ed. Athena Scientific, 1999. pages 24

[4] Brahme, A. Optimization of stationary and moving beam radiation therapy techniques. *Radiotherapy and Oncology 12* (1988), 129 – 140. pages 17

[5] Breedveld, S.; Bas van den, B., and Heijmen, B. An interior-point implementation developed and tuned for radiation therapy treatment planning. *Computational Optimization and Applications*, 68 (2017), 209–242. pages 18, 37, 53

[6] Breedveld, S., and Heijmen, B. Data for TROTS - the radiotherapy optimisation test set. *Data in Brief 12* (2017), 143–149. pages 20, 50, 53, 54, 66, 67

[7] Breedveld, S., and Heijmen, B. *TROTS - The Radiotherapy Optimisation TestSet*. https://www.erasmusmc.nl/en/cancer-institute/research/projects/radiotherapy-trots, 2019. pages 12, 50, 51, 53, 55, 57

[8] Breedveld, S., Storchi, P. R., Voet, P. W., and Heijmen, B. iCycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Medical Physics 39*, 2 (2012), 951–963. pages 12, 18, 54, 57

[9] Censor, Y., Altschuler, M., and Powlis, W. A computational solution of the inverse problem in radiation therapy treatment planning. *Applied Mathematics and Computation 25* (1988), 57–87. pages 38

[10] Cormack, A., and Quinto, E. The mathematics and physics of radiation dose planning using x-rays. *Contemporary Mathematics 113* (1990), 41–55. pages 38

[11] Dantzig, G. B. Programming of interdependent activities: II mathematical models. *Econometrica 17*, 3/4 (1949), 200–211. pages 16

[12] Drzymala, R., Mohan, R., Brewster, L., Chu, J., Goitein, M., Harms, W., and Urie, M. Dose-volume histograms. *International Journal of Radiation Oncology, Biology, Physics 21*, 1 (1991), 71–78. pages 56

[13] Ehrgott, M., Güler, C., Hamacher, H. W., and Shao, L. Mathematical optimization in intensity modulated radiation therapy. *Annals of Operations Research 175* (2010), 309 – 365. pages 18

[14] Holder, A. Designing radiotherapy plans with elastic constraints and interior point methods. *Health Care Management Science 6* (2003), 5 – 16. pages 18

[15] Iusem, A. Linear and nonlinear programming. In *International Encyclopedia of the Social Behavioral Sciences*, N. J. Smelser and P. B. Baltes, Eds. Pergamon, Oxford, 2001, pp. 8868–8874. pages 16

[16] Jamison, K. D., and Lodwick, W. A. Fuzzy linear programming using penalty methods. *Fuzzy Sets and Systems 119*, 1 (2001), 97 – 110. pages 19, 33

[17] Karmarkar, N. A new polynomial-time algorithm for linear programming. *Combinatorica 4* (1984), 373–395. pages 16

[18] Lodwick, W. A., and Bachman, K. A. Solving large-scale fuzzy and possibilistic optimization problems. *Fuzzy Optimization and Decision Making 4* (2005), 257 – 278. pages 19, 33

[19] Lodwick, W. A., Neumaier, A., and Newman, F. Optimization under uncertainty methods and applications in radiation therapy. *Proceedings 10th IEEE International Conference on Fuzzy Systems 3* (2001), 1219–1222. pages 19

[20] Neumaier, A. Fuzzy modeling in terms of surprise. *Fuzzy Sets and Systems 1*, 135 (2003), 21–38. pages 19, 33, 34, 39

[21] Nocedal, J., and Wright, S. J. *Numerical Optimization*. Springer, New York, NY, 2006. pages 19, 22

[22] Shepard, D. M., Ferris, M. C., Olivera, G. H., and Mackie, T. R. Optimizing the delivery of radiation therapy to cancer patients. *Society for Industrial and Applied Mathematics 41*, 4 (1999), 721 – 744. pages 18

[23] Tanaka, H., and Assai, K. Fuzzy linear programming based on fuzzy functions. *IFAC Proceedings Volumes 14*, 2 (1981), 785 – 790. pages 33

[24] Tanaka, H., and Assai, K. Fuzzy linear programming problems with fuzzy numbers. *Fuzzy Sets and Systems 13*, 1 (1984), 1 – 10. pages 33

[25] Tanaka, H., Okuda, T., and Assai, K. On fuzzy mathematical programming. *Journal of Cibernetics 3* (1974), 37 – 46. pages 19

[26] ur Rehman, J., Zahra, Ahmad, N., Khalid, M., ul Huda Khan Asghar, H. N., Gilani, Z. A., Ullah, I., Nasar, G., Akhtar, M. M., and Usmani, M. N. Intensity modulated radiation therapy: A review of current practice and future outlooks. *Journal of Radiation Research and Applied Sciences 11*, 4 (2018), 361–367. pages 17

[27] Werro, N. *Fuzzy Classification of Online Customers*. Fuzzy Management Methods. Springer International Publishing, 2015. pages 30, 31

[28] Wright, S. J. *Primal-Dual Interior-Point Methods*. Society for Industrial and Applied Mathematics, 1997. pages 16, 22, 24, 25, 26

[29] Zadeh, L. A. Fuzzy sets. *Information and Control 8*, 3 (1965), 338 – 353. pages 30

[30] Zimmermann, H.-J. Description and optimization of fuzzy systems. *International Journal of General Systems 2* (1976), 209–215. pages 19

[31] Zimmermann, H. J. Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems 1*, 1 (1978), 45 – 55. pages 33

[32] Zimmermann, H. J. Applications of fuzzy set theory to mathematical programming. *Information Sciences 36*, 1 (1985), 29 – 58. pages 33